


```

source      1. 1463
/organism="unknown"
BASE COUNT  466 a 292 c 361 g 344 t
ORIGIN

Query Match      100.0%; Score 1463; DB 6; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACGAGAAATAGTCGCGCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCCGCCCA 60
   |||||||
Db 1 CACGAGAAATAGTCGCGCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCCGCCCA 60

QY 61 CCCGCCACAGCGCGCAGAGATGGGATCAAGGGTTTCAGAAACTGCTGGCGGAC 120
   |||||||
Db 61 CCCGCCACAGCGCGCAGAGATGGGATCAAGGGTTTCAGAAACTGCTGGCGGAC 120

QY 121 AATGCGCCCAAGGGGATGAAGGAGCAGAAATTCGAGAGCTACTTCGGCGCGCAAAATCGCC 180
   |||||||
Db 121 AATGCGCCCAAGGGGATGAAGGAGCAGAAATTCGAGAGCTACTTCGGCGCGCAAAATCGCC 180

QY 181 GTCGACGCCAGCATGAGCATATACAGTCTCTGATTGTGAGTGGAAAGCAGAGCATGAA 240
   |||||||
Db 181 GTCGACGCCAGCATGAGCATATACAGTCTCTGATTGTGAGTGGAAAGCAGAGCATGAA 240

QY 241 ACTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTTTCGAGGAATGTTCAACCGGACA 300
   |||||||
Db 241 ACTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTTTCGAGGAATGTTCAACCGGACA 300

QY 301 ATAAGATTACTGGAAGGGGAATCAAGCCAGTTTATGTTTGTGTTGCAAGCCCTCTGAT 360
   |||||||
Db 301 ATAAGATTACTGGAAGGGGAATCAAGCCAGTTTATGTTTGTGTTGCAAGCCCTCTGAT 360

QY 361 ATGAAGAAACAGAGCTTGCTAAAGATCTCAAAAAGAGATGATCAACCAAGATCTG 420
   |||||||
Db 361 ATGAAGAAACAGAGCTTGCTAAAGATCTCAAAAAGAGATGATCAACCAAGATCTG 420

QY 421 ACTGAGCAGTAGAGTAGAGATGAAGATCGGATTCGAAATGAGCAGAGACTGTA 480
   |||||||
Db 421 ACTGAGCAGTAGAGTAGAGATGAAGATCGGATTCGAAATGAGCAGAGACTGTA 480

QY 481 AAGTCAAGAGGCAACACAGAGATTGTAACGGCTATTAGACTTATGGGGTTCCCT 540
   |||||||
Db 481 AAGTCAAGAGGCAACACAGAGATTGTAACGGCTATTAGACTTATGGGGTTCCCT 540

QY 541 GTTGTAGAGCACCCTCTGAGCAGAGAGAGATGTCGACCCCTTTCATTAACGATAAG 600
   |||||||
Db 541 GTTGTAGAGCACCCTCTGAGCAGAGAGAGATGTCGACCCCTTTCATTAACGATAAG 600

QY 601 GTGTTCTGCTGTGCTTCAGAGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCCT 660
   |||||||
Db 601 GTGTTCTGCTGTGCTTCAGAGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCCT 660

QY 661 CGTCATTAATGATGATCCCAAGAAATACCTGTGATGAAATTTGATGTTGCAAG 720
   |||||||
Db 661 CGTCATTAATGATGATCCCAAGAAATACCTGTGATGAAATTTGATGTTGCAAG 720

QY 721 GTTTTGGAGGAGCTTGAACATCACCATGGACAGTTCATTGATTTGTCATCTGTGGA 780
   |||||||
Db 721 GTTTTGGAGGAGCTTGAACATCACCATGGACAGTTCATTGATTTGTCATCTGTGGA 780

QY 781 TGTGACTATTGTGATGACATCAAGGATATCGGGGGCAACACACTCTGAAACTTATTGCT 840
   |||||||
Db 781 TGTGACTATTGTGATGACATCAAGGATATCGGGGGCAACACACTCTGAAACTTATTGCT 840

QY 841 CAACATGGGTCCATAGAACATCTTTGGAGATCTTTAATAAGACAGATATCAAAATTCCT 900
   |||||||
Db 841 CAACATGGGTCCATAGAACATCTTTGGAGATCTTTAATAAGACAGATATCAAAATTCCT 900

QY 901 GAGGACTGGCCTTACCAAGAGCTGCACGCTTGTTCAGGAGCCTAATGTCACATTGGAT 960
   |||||||
Db 901 GAGGACTGGCCTTACCAAGAGCTGCACGCTTGTTCAGGAGCCTAATGTCACATTGGAT 960

```

```

QY 961 ATTCTGAGCTAAATGAGTGCACCTGATGAGGAGGCTCTCATAGTTTCTGCTGTA 1020
   |||||||
Db 961 ATTCTGAGCTAAATGAGTGCACCTGATGAGGAGGCTCTCATAGTTTCTGCTGTA 1020

QY 1021 GATAATGTTTCAACGAGATCGGTTGACAAAGGCCATAGAGAGATCAATCTGCCAAG 1080
   |||||||
Db 1021 GATAATGTTTCAACGAGATCGGTTGACAAAGGCCATAGAGAGATCAATCTGCCAAG 1080

QY 1081 AATAAATCTCGCAAGGAAGACTCGAGTCCCTTTTCAAGCCCAACTGCCACACATCAGCA 1140
   |||||||
Db 1081 AATAAATCTCGCAAGGAAGACTCGAGTCCCTTTTCAAGCCCAACTGCCACACATCAGCA 1140

QY 1141 CCGTAAACCGAAGGAGACTTCGGATATAAACAAGCAAGGAGCTGCCGAACAAGAAACA 1200
   |||||||
Db 1141 CCGTAAACCGAAGGAGACTTCGGATATAAACAAGCAAGGAGCTGCCGAACAAGAAACA 1200

QY 1201 AAGCTGGTGAAGAGAAATAATCTTGGATGCTTGTATGATACAACTACGACTACGAAAG 1260
   |||||||
Db 1201 AAGCTGGTGAAGAGAAATAATCTTGGATGCTTGTATGATACAACTACGACTACGAAAG 1260

QY 1261 CAGCGGTGGCTGATCACTTCGCTTAGATTTAATTAACCTCCCTGTTTAACTCAGAGCTTT 1320
   |||||||
Db 1261 CAGCGGTGGCTGATCACTTCGCTTAGATTTAATTAACCTCCCTGTTTAACTCAGAGCTTT 1320

QY 1321 GGTAAAAGTTTGGCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTTGAAGAGATTGCTG 1380
   |||||||
Db 1321 GGTAAAAGTTTGGCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTTGAAGAGATTGCTG 1380

QY 1381 TACCAAGTAAACAACATATCGCTGTTTCTTACTTCTTGCTTTGAAGTAAACAAAAA 1440
   |||||||
Db 1381 TACCAAGTAAACAACATATCGCTGTTTCTTACTTCTTGCTTTGAAGTAAACAAAAA 1440

QY 1441 AAAAAA 1463
Db 1441 AAAAAA 1463

```

```

RESULT 2
AR152406 LOCUS 1478 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 7 from patent US 6232527.
ACCESSION AR152406
VERSION AR152406.1 GI:15118456
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1478)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 7 15-MAY-2001;
FEATURES
    Location/Qualifiers
        source
            1..1478
            /organism="unknown"
BASE COUNT 463 a 302 c 365 g 348 t
ORIGIN

```

```

Query Match      97.2%; Score 1421.8; DB 6; Length 1478;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 AATAGCTCGGGCGCGGTTCTTTCGCCACATCCGGCTCAGCGCGCGCCGCCCGCC 66
   |||||||
Db 19 AATAGCTCGGGCGCGGTTCTTTCGCCACATCCGGCTCAGCGCGCGCCGCCCGCC 78

QY 67 ACAGCGCGCGCAGACGAGATGGGCATCAAGGGTTTCAGCAAACTGCTGGCGACAATGG 126
   |||||||
Db 79 ACAGCGCGCGCAGACGAGATGGGCATCAAGGGTTTCAGCAAACTGCTGGCGACAATGG 138

QY 127 CCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGGCAAAATCGCGTCGAC 186
   |||||||
Db 139 CCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGGCAAAATCGCGTCGAC 198

```

QY	187	GCAGCATGAGCATATACCAAGTTCCTGATGTAGTTGGAAGGACAGCATGGAAACTCTC	246
Db	199	GCCAGCATGAGCATCTACCAAGTTCCTGATAGTAGTTTGGAAAGGACAGCATGGAAACTCTC	258
QY	247	ACAAATGAAGCTGGTGAAGTCACCTAGTTCATTTGCAAGGAATGTTCAACCGGACAATAAGA	306
Db	259	ACAAATGAAGCTGGTGAAGTTCACCTAGTTCATTTGCAAGGAATGTTCAACCGGACAATAAGA	318
QY	307	TTACTGGAAGCGGGAAATCAAGCCAGTTTATGTTTGTATGGCAAGCTCCTGATATGAAG	366
Db	319	TTACTGGAAGCGGGAAATCAAGCCAGTTTATGTTTGTATGGCAAGCTCCTGATATGAAG	378
QY	367	AAACRAGAGCTTGCCTAAAGATACCTCAAAAGAGATGATCAACCAAGAGATCTGACTGAG	426
Db	379	AAACRAGAGCTTGCCTAAAGATACCTCAAAAGAGATGATCAACCAAGAGATCTGACTGAG	438
QY	427	GCAGTAGAGGTAGGAGATAAAGATGCGATTGAAAATTTGACGAAGAGGACTGTAAAGGTC	486
Db	439	GCAGTAGAGGTAGGAGATAAAGATGCGATTGAAAATTTGACGAAGAGGACTGTAAAGGTC	498
QY	487	ACAAGGCACACACGAAGATTGTAACCGCTATTAGACTTATGGGGGTTCCGTGTTGTA	546
Db	499	ACAAGGCACACACGAAGATTGTAACCGCTATTAGACTTATGGGGGTTCCGTGTTGTA	558
QY	547	GAGGCACCTTCTGAAGCAGAAGCAGAATGTGCAGCCCTTTGCATAACGATAAGGTGTTTC	606
Db	559	GAGGCACCTTCTGAAGCAGAAGCAGAATGTGCAGCCCTTTGCATAACGATAAGGTGTTTC	618
QY	607	GCTGTTGCTTCAGAAGATATGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCAT	666
Db	619	GCTGTTGCTTCAGAAGATATGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCAT	678
QY	667	TTAATGGATCCAAGTTCCAAAGAAAATACCTGTATGGAATTTGATGTTGCCAAGGTTTGG	726
Db	679	TTAATGGATCCAAGTTCCAAAGAAAATACCTGTATGGAATTTGATGTTGCCAAGGTTTGG	738
QY	727	GAGGAGCTTGAACCTCACCATGGACCAAGTTTCATTTGATTTGTCATCCTGTGTGGATGTGAC	786
Db	739	GAGGAGCTTGAACCTCACCATGGACCAAGTTTCATTTGATTTGTCATCCTGTGTGGATGTGAC	798
QY	787	TATTTGATAGCATCAAGGATTCGGGGGCAACACAGCTCTGAAACTTATTCTGCTCAACAT	846
Db	799	TATTTGATAGCATCAAGGATTCGGGGGCAACACAGCTCTGAAACTTATTCTGCTCAACAT	858
QY	847	GGGTCCATAGAAAAGCATCTTGGAGATCTTATAAAGACAGATATCAAAATTCCTGAGGAC	906
Db	859	GGGTCCATAGAAAAGCATCTTGGAGATCTTATAAAGACAGATATCAAAATTCCTGAGGAC	918
QY	907	TGGCCTTACCAAGACCTCGAGCGTTTGTCAAGGAGCCCTAATGTTCACATTTGGATATTCCT	966
Db	919	TGGCCTTACCAAGACCTCGAGCGTTTGTCAAGGAGCCCTAATGTTCACATTTGGATATTCCT	978
QY	967	GAGCTAAATTTGAGCTGCACCTGATGAGGAGGCTTCATTAAGTTTCCCTGGTAAAGATTAAT	1026
Db	979	GAGCTAAATTTGAGCTGCACCTGATGAGGAGGCTTCATTAAGTTTCCCTGGTAAAGATTAAT	1038
QY	1027	GTTTTCACGAGATCGGTTGACAAAGGCCATAGAGAGATCAAAATCTGCCAAGAATAAA	1086
Db	1039	GTTTTCACGAGATCGGTTGACAAAGGCCATAGAGAGATCAAAATCTGCCAAGAATAAA	1098
QY	1087	TCGTGCAAGGACACTCGAGTCTCTTTTTCAAAGCAACTGGCCACCACATCAGCACCGCTA	1146
Db	1099	TCGTGCAAGGACACTCGAGTCTCTTTTTCAAAGCAACTGGCCACCACATCAGCACCGCTA	1158
QY	1147	AAACGGAGGAGACTTCGGGATAAAACAAGGACGTCGGGACAGAAAACAAGGCT	1206
Db	1159	AAACGGAGGAGACTTCGGGATAAAACAAGGACGTCGGGACAGAAAACAAGGCT	1218
QY	1207	GGTGGAAGAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAACACGCGG	1266
Db	1219	GGTGGAAGAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAACACGCGG	1278
QY	1267	TGGGGTGATCACTTCGGTCTAGATTAATTTAACTCCCTGTTTAACTCAGACGCTTTGGTAAA	1326

Db	1279	TGGCATGATCACTTCGCGCTAGATTATTTAACTCCCTGTTTAACTCAGACCTTTGGTGAA	1338
Qy	1327	AGTTTGGCTCATGTTTCAAGCTGGGTAAGTTAGTTGTCTTTTCAAGAGAGATTGGTGATACAA	1386
Db	1339	AGTTTGGCCATGTTTCAAGCTGGGTAAGTTAGTTGTCTTTGAAGAGAGATTGGTGATACAA	1398
Qy	1387	GTAACAAACACTATCGCTGTGTTTTTACCTCTGTCCTTGAAGTAAAAAATAAAAAA	1446
Db	1399	GTAACAAACACTATCGCTGTGTTTTTACCTCTGTCCTTGAAGTATGATCCAGTAAAA	1458
Qy	1447	AAAAAATAAAAAAATAAAAAA 1463	
Db	1459	AAAAAATAAAAAAATAAAAAA 1475	
RESULT 3			
LOCUS	AR152404	1541 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 3 from patent US 6232527.		
ACCESSION	AR152404		
VERSION	AR152404.1	GI:15118454	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1541)		
TITLE	Mahajan, P.B.		
JOURNAL	Maize Rad2/FEN-1 orthologues and uses thereof		
FEATURES	Patent: US 6232527-A 3 15-MAY-2001; Location/Qualifiers 1..1541 /organism="unknown"	383 t	
BASE COUNT	473 a 308 c 377 g		
ORIGIN			
	Query Match	96.4%;	Score 1409.8; DB 6; Length 1541;
	Best Local Similarity	99.9%;	Pred. No. 0;
	Matches 1411; Conservative	0;	Mismatches 2; Indels 0; Gaps 0
Qy	19	CGCGCGTTCTTCGCGCACTTCGGCTCAGCGCGCGCCGCGCCACCGCCACAGCC	78
Db	13	CGCGCGTTCTTCGCGCACTTCGGCTCAGCGCGCGCCGCGCCACCGCCACAGCC	72
Qy	79	GACGAGATGGCATCAAGGTTTGACAAACTGCTGGCGGACAAATGCGCCAGGCGATG	138
Db	73	GACGAGATGGCATCAAGGTTTGACAAACTGCTGGCGGACAAATGCGCCAGGCGATG	132
Qy	139	AAGGACGAGAGTTCAGAGCTACTTCGCGCGCAAAATCGCGTCGACGCCAGCATGAGC	198
Db	133	AAGGACGAGAGTTCAGAGCTACTTCGCGCGCAAAATCGCGTCGACGCCAGCATGAGC	192
Qy	199	ATATACAGTTCCTGATTGTAGTTGGAAGGACAGGCGATGGAACCTCTCAAAATGAAGCT	258
Db	193	ATATACAGTTCCTGATTGTAGTTGGAAGGACAGGCGATGGAACCTCTCAAAATGAAGCT	252
Qy	259	GGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGCACATAGATTACTTGGAGCG	318
Db	253	GGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGCACATAGATTACTTGGAGCG	312
Qy	319	GGAATCAAGCGCTTATGTTTGTATGGCAAGCTTCCTGATATGAAGAAACAAGAGCTT	378
Db	313	GGAATCAAGCGCTTATGTTTGTATGGCAAGCTTCCTGATATGAAGAAACAAGAGCTT	372
Qy	379	GCTAAAAGATACATAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGATGAGGTA	438
Db	373	GCTAAAAGATACATAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGATGAGGTA	432
Qy	439	GGAGATAAAGATCGGATTGAAAATTCAGCAAGAGGACTGTAAGGTCACAAAGGCAACAC	498
Db	433	GGAGATAAAGATCGGATTGAAAATTCAGCAAGAGGACTGTAAGGTCACAAAGGCAACAC	492
Qy	499	AAGCAAGATTGTAAACGGCTATTAAAGACTTATGGGGTTCCTCTGTTAGAGGCACCTTCT	558

```

Db 493 AACGAGATTGTAACGGCTATTAGACTATGGGGTTCCTGTTGTAGAGCACCTTCT 552
QY 559 GAAGCAGAGCAGATGTGACCCCTTGGCATAAACGATAAGGTGTGCTGCTTCTTCA 618
Db 553 GAAGCAGAGCAGATGTGACCCCTTGGCATAAACGATAAGGTGTGCTGCTTCTTCA 612
QY 619 GAAGATATGGACTCCCTTACTTCTTGGGGTCCACGGTTCCTTCGTCATTTATGGATCCA 678
Db 613 GAAGATATGGACTCCCTTACTTCTTGGGGTCCACGGTTCCTTCGTCATTTATGGATCCA 672
QY 679 AGTTCCTCAAGAAATACCTGATGAAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGA 738
Db 673 AGTTCCTCAAGAAATACCTGATGAAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGA 732
QY 739 CTCACCATGGACAGTTCATTGATTTGTCATCTCTGTCGATGTCATTTATGGATAGC 798
Db 733 CTCACCATGGACAGTTCATTGATTTGTCATCTCTGTCGATGTCATTTATGGATAGC 792
QY 799 ATCAAGATATCGGGGGGCAACAGCTCTGAACCTTATTCGTCACATGGGTTCATAGAA 858
Db 793 ATCAAGATATCGGGGGGCAACAGCTCTGAACCTTATTCGTCACATGGGTTCATAGAA 852
QY 859 AGCATCTTTGGAGAACTTAAATAAGACAGATCAAAATTCCTGAGCTGGCCTTACCAA 918
Db 853 AGCATCTTTGGAGAACTTAAATAAGACAGATCAAAATTCCTGAGCTGGCCTTACCAA 912
QY 919 GAAGCTGCACCTTGTTCAGAGGCTTAATGTCACATTTGATATTCCTGAGCTAAAATGG 978
Db 913 GAAGCTGCACCTTGTTCAGAGGCTTAATGTCACATTTGATATTCCTGAGCTAAAATGG 972
QY 979 ACTGCACCTGATGAGGAGGTCTCATAGTTTCTGGTAAAGATATATGTTTCAACGAA 1038
Db 973 ACTGCACCTGATGAGGAGGTCTCATAGTTTCTGGTAAAGATATATGTTTCAACGAA 1032
QY 1039 GATCGGTGACAAAGGGCCATAGAAGATCAAAATCTGCAAGAAATAATCTGCAAGGA 1098
Db 1033 GATCGGTGACAAAGGGCCATAGAAGATCAAAATCTGCAAGAAATAATCTGCAAGGA 1092
QY 1099 AGACTCGAGTCTTTTCAAGCCAACTGCCACACATCAGACCGCTAAAACGGAAGAG 1158
Db 1093 AGACTCGAGTCTTTTCAAGCCAACTGCCACACATCAGACCGCTAAAACGGAAGAG 1152
QY 1159 ACTTCGGATAAAACAAGCGAGCTGCGAACAAGAAACAAGGCTGTGTTGAAAGAG 1218
Db 1153 ACTTCGGATAAAACAAGCGAGCTGCGAACAAGAAACAAGGCTGTGTTGAAAGAG 1212
QY 1219 AAATAATCTTGGATGCTGTATGTAACAACCTAGCACTAGCAAGAGCGTGGCGTGTATCAC 1278
Db 1213 AAATAATCTTGGATGCTGTATGTAACAACCTAGCACTAGCAAGAGCGTGGCGTGTATCAC 1272
QY 1279 TTGCTTTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTCTCATG 1338
Db 1273 TTGCTTTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTCTCATG 1332
QY 1339 TTTCAAGCTGGGTAACTTGTGTTGTTGAAGAGATTGTTGACCAAGTAACAAACTT 1398
Db 1333 TTTCAAGCTGGGTAACTTGTGTTGTTGAAGAGATTGTTGACCAAGTAACAAACTT 1392
QY 1399 ATCGCTGTTTTTACTTCTTGTCTTTGAAAGTA 1431
Db 1393 ATCGCTGTTTTTACTTCTTGTCTTTGAAAGTA 1425

```

```

RESULT 4
AR152405
LOCUS AR152405 1381 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6232527.
ACCESSION AR152405
VERSION AR152405.1 GI:15118455
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```

```

Unclassified.
REFERENCE 1 (bases 1 to 1381)
AUTHORS Mahajan, P. B.
TITLE Maize Rad2/FBN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 5 15-MAY-2001;
FEATURES Location/Qualifiers
source 1. 1381
BASE COUNT 441 a 269 c 346 g 325 t
ORIGIN
Query Match 92.3%; Score 1350.2; DB 6; Length 1381;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 49 CGCCGCCGCCACCGCCACACGCCGCCACGAGATGGGATCGGATCAAGGGTTTGACGAAA 108
Db 1 CGACCCACCGCTCCGCGCCACAGCCGCCACGAGATGGGATCGGATCAAGGGTTTGACGAAA 60
QY 109 CTGCTGGCGGACAATGCGCCCAAGCGGATGAAGAGCAGAAAGTTTCGAGAGCTACTTCGCG 168
Db 61 CTGCTGGCGGACAATGCGCCCAAGCGGATGAAGAGCAGAAAGTTTCGAGAGCTACTTCGCG 120
QY 169 CGCAAAATCGCCGTCGCGAGCCAGCATATACCAAGTTTCTGATGTAGTTGGAAAG 228
Db 121 CGCAAAATCGCCGTCGCGAGCCAGCATATACCAAGTTTCTGATGTAGTTGGAAAG 180
QY 229 ACAGGCTATGAAACTCTCAAAATGAAGTCTGTAAGTCACTAGTCAATTTGCAAGGAATG 288
Db 181 ACAGGCTATGAAACTCTCAAAATGAAGTCTGTAAGTCACTAGTCAATTTGCAAGGAATG 240
QY 289 TTCAACCGGACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGC 348
Db 241 TTCAACCGGACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGC 300
QY 349 AAGCCTCTGTATGAGAAACAAGAGCTTCTTAAAGATACTCAAAAGAGATGATGCA 408
Db 301 AAGCCTCTGTATGAGAAACAAGAGCTTCTTAAAGATACTCAAAAGAGATGATGCA 360
QY 409 ACCAAGATCTGACTGAGGCACTAGAGTAGGATGAAGATGCGATTGAAAAATTTGAGC 468
Db 361 ACCAAGATCTGACTGAGGCACTAGAGTAGGATGAAGATGCGATTGAAAAATTTGAGC 420
QY 469 AAGAGGACTGTAAAGGTCACAAGGCAACAACAAGAGATTGTAAACGGCTATTAAAGACTT 528
Db 421 AAGAGGACTGTAAAGGTCACAAGGCAACAACAAGAGATTGTAAACGGCTATTAAAGACTT 480
QY 529 ATGGGGTTCCTGTTGTAGAGGCACTTCTCAAGCAGAGAGATGTGACGCCCTTTCG 588
Db 481 ATGGGGTTCCTGTTGTAGAGGCACTTCTCAAGCAGAGAGATGTGACGCCCTTTCG 540
QY 589 ATAAACGATAAGGTGTTTCGCTGTTTCAGAGATATGGACTCCCTTACTTTTGGGGCT 648
Db 541 ATAAACGATAAGGTGTTTCGCTGTTTCAGAGATATGGACTCCCTTACTTTTGGGGCT 600
QY 649 CCAGGTTTCTTCTGTCATTTTAAATGGATCAAGTTTCCAAAGAAATACCTGTGATGGAATTT 708
Db 601 CCAGGTTTCTTCTGTCATTTTAAATGGATCAAGTTTCCAAAGAAATACCTGTGATGGAATTT 660
QY 709 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCACCATGACCAAGTTCATTTGTTGTCG 768
Db 661 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCACCATGACCAAGTTCATTTGTTGTCG 720
QY 769 ATCCTGTGTGATGTGACTATTTGTATAGATCAAAAGTATCGGGGGGCAACAGCTCTG 828
Db 721 ATCCTGTGTGATGTGACTATTTGTATAGATCAAAAGTATCGGGGGGCAACAGCTCTG 780
QY 829 AAACCTTATCGTCAACATGGTCCATAGAAAGCATCTTGAGAGATCTTATAAAGACAGA 888
Db 781 AAACCTTATCGTCAACATGGTCCATAGAAAGCATCTTGAGAGATCTTATAAAGACAGA 840
QY 889 TATCAAAATCTCGAGGACTGGCCTTACCAGAAAGCTTCGAGCTTGTTCAGAGGAGCCTAAT 948

```



```
Db 841 TATCAAAATTCCTGAGGACTGGCCTTACCAAGAAGCTCGAGCGTTGTTCAAGGAGCCTAAT 900
QY 949 GTCACATTTGATATTCCTGAGCTAAATGAGCTGCACCTGATGAGGAGGTCTCATAAGT 1008
Db 901 GTCACATTTGATATTCCTGAGCTAAATGAGCTGCACCTGATGAGGAGGTCTCATAAGT 960
QY 1009 TTCCTGGTAAAGATAATGGTTTCAACGAGAGATCGGTCGACAAAGCCATAGAGAAGATC 1068
Db 961 TTCCTGGTAAAGATAATGGTTTCAACGAGAGATCGGTCGACAAAGCCATAGAGAAGATC 1020
QY 1069 AAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAGTCCCTTTTTCACAGCCAACTGCC 1128
Db 1021 AAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAGTCCCTTTTTCACAGCCAACTGCC 1080
QY 1129 ACCACATCAGCACCGCTAAACCGAAGGAGACTTCGGATAAACCAAGCAAGGAGCTGGC 1188
Db 1081 ACCACATCAGCACCGCTAAACCGAAGGAGACTTCGGATAAACCAAGCAAGGAGCTGGC 1140
QY 1189 AACAAAGAAACAAGGCTGTGGAAGAAGAAATAATCTTGGATGCTTGATGTACAACTA 1248
Db 1141 AACAAAGAAACAAGGCTGTGGAAGAAGAAATAATCTTGGATGCTTGATGTACAACTA 1200
QY 1249 CGACTACGAAGCAGCGGTGGCTGATCACTTCGCTTAGATTAATTTAACTCCCTGTTTA 1308
Db 1201 CGACTACGAAGCAGCGGTGGCTGATCACTTCGCTTAGATTAATTTAACTCCCTGTTTA 1260
QY 1309 ACTCAGAGCTTTGGTAAAGTTTGCTCATGTTTCAAGCTGGGTAAAGTTAGTTGTGTTG 1368
Db 1261 ACTCAGAGCTTTGGTAAAGTTTGCTCATGTTTCAAGCTGGGTAAAGTTAGTTGTGTTG 1320
QY 1369 AAGAGATTGCTGTACCAAGTAAACAAAACCTTATCGCTCTTTTTT 1411
Db 1321 AAGAGATTGCTGTACCAAGTAAACAAAACCTTATCGCTCTTTTTT 1363

RESULT 5
AB021666 1354 bp mRNA linear PLN 19-JUL-2002
LOCUS Oryza sativa (japonica cultivar-group) OSFEN-la mRNA for FEN-1,
DEFINITION complete cds.
ACCESSION AB021666
VERSION AB021666.2 GI:4587224
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare)
meristem cDNA to mRNA.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and
Sakaguchi,K.
TITLE Plant homologue of flap endonuclease-1: molecular cloning,
characterization, and evidence of expression in meristematic
tissues
JOURNAL Plant Mol. Biol. 42 (3), 415-427 (2000)
MEDLINE 20256470
REFERENCE 2 (bases 1 to 1354)
AUTHORS Kimura,S., Hashimoto,J. and Sakaguchi,K.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1998) Seisuke Kimura, Science University of
Tokyo, Dept. of Applied Biological Science; 2641 Yamazaki, Noda,
Chiba 278-8510, Japan (E-mail:seisuke@rs.noda.tus.ac.jp,
Tel:81-471-24-1501(ex.3409), Fax:81-471-23-9767)
COMMENT On Apr 17, 1999 this sequence version replaced gi:4062866.
FEATURES
source
1. .1354
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:3994"
/tissue_type="meristem"
1. .1354
/gene="OSFEN-la"
gene
```

```
CDS
67. .1209
/gene="OSFEN-la"
/function="endonuclease"
/codon_start=1
/product="FEN-1"
/protein_id="BAA36171.1"
/db_xref="GI:4587225"
/translation="MGIKGLTKLLADNAPKAMKEQKESYFGRRIVADSMSTYQFLI
VYVGTGMEITLNEAGEVTSHQGMENRTIRLLEAGIKPVYVDFDKPDLKQKELAKRY
SKREDATKELTEAVEEGDKDAIEKESKRVKVTQKHNECKRLRLMLGVVYVVEACEA
EABCAALCINDMVIYAVASDMDSLTGAPRFLHLMDSPKKI.PVMEFVAVKYLEELE
LTMDOFIDLCISGGDCDSIKIGGGTALKLIRHSGSLESLINKRQYQIPEDMP
YQEARLFKEPNVTLDIPELKNWAPDEGLVEFLVKEGFMNODRVTKAIEKIKFAPKNK
SSQRLSEFFKPVWTSVPLKRKDTSEKPTKAVANKTKGAGGKKF"
polyA_site
1354
/genes="OSFEN-la"
/note="17 A nucleotides"
BASE COUNT 415 a 286 c 331 g 322 t
ORIGIN
Query Match 63.4%; Score 927.4; DB 8; Length 1354;
Best Local Similarity 86.3%; Pred. No. 3.4e-221;
Matches 1037; Conservative 0; Mismatches 161; Indels 3; Gaps 1;
QY 55 CCGCCACCCCGCCAGCGCCGAGAGATGGGCATCAAGGGTTTGACGAAACTGCTG 114
Db 37 CGCGCGCTCGCGCATCCGTCGGAGGCGAGATGGGCATCAAGGGTTTGACGAACTTTG 96
QY 115 GCGGACAATGCGCCCAAGGGGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCCGCAAA 174
Db 97 GCGGACAATGCGCCCAAGGGGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCCGCGC 156
QY 175 ATCCCGCTGCAGCGCCAGCATGAGCATATACCAGTTCCTCTGTTGTAAGGACAGGC 234
Db 157 ATTCCCGCTGCAGCGCCAGCATGAGCATCTCCAGTTCCTCTGTTGTAAGGACAGGC 216
QY 235 ATGAAACTCTCACAATAAGCTGGTGAAGTCACTAGTCACTATTCGCAAGGAATGTTCAAC 294
Db 217 ATGAAACCCTCAGCAAGGAGCGCGGTGAAGTCACCAAGTATGTTGCAAGGTATGTTCAAC 276
QY 295 CGGACAATAGATTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTGTGAGGAAAGCCT 354
Db 277 CGGACAATAAGTTGCTGGAGGAGGAATCAAGCCAGTATATGTTTGTGAGCGGAAGCCT 336
QY 355 CCTGATATGAAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAACCAAAA 414
Db 337 CCTGATCTGAAGAAACAAGAGCTCGCAAAAGAGTACTCGAAGAGGAGGATGCCACCAAA 396
QY 415 GATCTGACTGAGGCGAGTAGAGGTAGAGATAAAGATGCGATTGAAAATTTGAGCAAGAGG 474
Db 397 GAACCTAACAGAGGCGAGTAGAGGAGGAGATAAAGACGCAATTTGAAAATTTGAGCAAGAG 456
QY 475 ACTGTAAGGTCACAAGGCAACACCAAGCATTTCTAAACGCTATTAAAGACTTATGCGG 534
Db 457 ACTGTGAAGGTCACGAAGCAGCAATGAAGAATGTAACGACTATTAAGACTTATGCGG 516
QY 535 GTTCCTGTTGTAGAGGCACTTCTGAAGCAGAAGCAGAAATGTGCGCCCTTTTGCATAAAC 594
Db 517 GTTCCTGTTGTAGAGGCGCCATGTGAACGAGCAGAATGTGCTGCACCTTTGCAATAAT 576
QY 595 GATAAGGTTTCGCTGTTGCTTTCAGAGATATGAGTCCCTTACTTTTGGGCGTCACGG 654
Db 577 GATATGTTTATGCAAGTTGTCATCAGAAGATGGAAGTCTCTTACTTTTGGAGCTCCAGG 636
QY 655 TTCCTCTCTCATTTAATGATCCCAAGTTCCCAAGAAATACCTGTGATGAA-TTTTATGTT 714
Db 637 TTCCTCTCTCATTTGATGGATCCCAAGTTCCCAAGAAATTTCCAGTTATGAAATTTGAATC 696
QY 715 GCCAAGGTTTGGAGGAGCTTGAACCTCACCAGTCCAGTTCATTGATTTTCGATCCTG 774
Db 697 GCGAAAGTTTGGAGGAGCTTGAACCTCAGTCTACTATGACCAATTCATTGCTGTCATCCTA 756
QY 775 TGTGATGTGACTATTGTGTAGTACATCAAGGATCAAGGATCGGGGGGCAACAGCTCTGAACTT 834
```

```

Db 757 TCTGGATGTCATATTTGTGATAGCATTAAGGTATTTGGAGTCAACAGCTCTGAAACTT 816
QY 835 ATTTCGTCAACATGGTCCATAGAACCATCTTTGGAGAACTTTAATAAAGACAGATATCAA 894
Db 817 ATTTCGTCAACATGGTCCATAGAACCATCTTGGAAACATTAATAAAGACAGATATCAG 876
QY 895 ATTTCGTGAGACTGGCTTTACCAAGAAGCTCGACGCTTGTTCAGAGAGCTTAATGTACCA 954
Db 877 ATTTCGTGAGACTGGCTTTATCAAGAAGCTCGACGCTTGTTCAGAGAACCCCAATGTTACA 936
QY 955 TTGGATATTCCTCAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAGTTTCCTG 1014
Db 937 TTGGATATTCCTCAGCTAAATGGAAATGCTCCCGATGAGGAAGGCTTGTAGAGTTTCTG 996
QY 1015 GTAAAAAATGTTTCAACGAAGACTCGGTGACAAAGGCCATAGAGAAGATCAAACT 1074
Db 997 GTCAAGAAATGTTTCAACCAAGATAGAGTGACAAAGGCCATAGAGAAGATTAATTT 1056
QY 1075 GCCAAGAAATGCTCGCAAGAGACTCGAGTCTCTTTTCAAGCCCACTGCCACCACA 1134
Db 1057 GCCAAGAAATGCTTCCCAAGGAAGGCTCGAATCCTCTTCAAGCCAGTTGTAGCACA 1116
QY 1135 TCAGCACCTTAACCGGAGAGACTTCGGATAAACAAGCAAGCAGCAGCTCCGACACAG 1194
Db 1117 TCAGTCCCTCTGAAGAAGAGACACTTCAGAAACCAACCAAGCAGCTCGCTAACCAAG 1176
QY 1195 AAAACAAG--GCTGGTGAAGAAGAAATAATCTTGATGCTTGTATCAACTACGA 1251
Db 1177 AAAACAAGGAGCGCGGCGAAGAAATAAGTTCGATTGTGCAATTCATCAACAAGA 1236
QY 1252 C 1252
Db 1237 C 1237

RESULT 6
XLU68141
LOCUS
DEFINITION
Xenopus laevis xFEN1b mRNA, complete cds.
ACCESSION
U68141
VERSION
U68141.1 GI:1549392
KEYWORDS
Xenopus laevis.
SOURCE
Xenopus laevis.
ORGANISM
Xenopus laevis.
REFERENCE
1 (bases 1 to 1469)
Bibikova, M., Wu, B., Chi, E., Kim, K.-H., Trautman, J. K. and Carroll, D.
Characterization of FEN-1 from Xenopus laevis. cDNA cloning and
role in DNA metabolism
J. Biol. Chem. 273 (51), 34222-34229 (1998)
MEDLINE
99069415
PUBMED
9852084
REFERENCE
2 (bases 1 to 1469)
Bibikova, M., Chi, E., Wu, B., Kim, K.-H. and Carroll, D.
Direct Submission
AUTHORS
Submitted (27-AUG-1996) Biochemistry, Univ. of Utah, 50 N. Medical
Drive, Salt Lake City, UT 84132, USA
JOURNAL
Location/Qualifiers
FEATURES
source
1. 1469
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="head"
/clone_lib="Hemmati-Briuanlou et al. (1991) Development
111, 715-724"
/dev_stage="tadpole"
/note="also present in oocytes, as determined by
sequencing of RT-PCR products"
120..1268
/function="5'-3' exonuclease and a structure-specific
endonuclease that removes 5' single-stranded 'flaps';

```

```

implicated in processing of Okazaki fragments during DNA
replication and in base excision repair; 5' nuclease"
/note="FEN-1 homolog; similar to mammalian Dnase IV or
MFL, to S. cerevisiae RAD27, RTH1 or YKL510, and to S.
pombe rad2; similar to Xenopus laevis xFEN1a encoded by
GenBank Accession Number U64563"
/codon_start=1
/product="xFEN1b"
/protein_id="AAB08478.1"
/db_xref="GI:1549393"
/translation="MGIHLAKLIADVAPAAIKHDIKSYFGRKVAVDASMCYIOFLI
AVRDGNMLQNEEGETTSHLMFKYIRLMEHGKIPVYVFDGPKQKSGSLAKRSE
RAAEAKLEEAEEAGEVENIEKFKRLVYKQHNECKKLLSLMGIPYVDPCEAE
ATCAALVKAGKYAAATDMDALTCTPVLRLHJASEAKLPIQEFHLNRYFDQIGI
NTEQFVDCILLGSDYCEITRIGIPKRAIDLJROHKTLEIINDIDNKKYPIPEWNLH
KEAROLFLEPEVIDADITELKWTEDDEGLVAFMGKEQFSEIRINRNGAKKLAKNRQG
STQGRLLDFFKVTGSSISSTKRKEVSGSKTKSKTKTGTPAGFKRKG"
BASE COUNT 461 a 288 c 377 g 343 t
ORIGIN
Query Match 22.3%; Score 326.8; DB 5; Length 1469;
Best Local Similarity 58.1%; Pred. No. 8.5e-71;
Matches 615; Conservative 0; Mismatches 437; Indels 6; Gaps 2;
QY 78 AGACGAGATGGGCATCAAGGGTTTTCGAAACTCTCGCGGACAATCGGCCCAAGCGAT 137
Db 113 AGTCAATATGGGAATTCACGGCTTGCCCAAGCTTATCGCAGATGTGCGCCTGCACCTAT 172
QY 138 GAAGGACGAGAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCGTCGACGCCAGATGAG 197
Db 173 CAAAGAGCATGATATAAAAGTTACTTTGGTCGAAAAGTGGCGGTTGATGCTTCTATGTG 232
QY 198 CATATACCACTTCTGATTTAGTTGGAAGGACAGGACGCTGGAACCTCTCACAAATCAAGC 257
Db 233 CATTTACCAGTTCCTTATCGCTGTGACAGAGGATGSCA---ATATGCTGCAGAAATGAGGA 289
QY 258 TGTGAAGTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACAATTAAGATTACTTGAAGC 317
Db 290 AGCGAAACACCAAGGCATCTCATGGGTATGTTTACCGTACCAATTCGCATGTTGAGCA 349
QY 318 GGAATCAAGCCAGTTTATGTTTTTTCGAAAGCTCTCTGATATGAACAAACAGAGCT 377
Db 350 TGGCATCAAAACCCGTTATGTTTATGGGAAGCCGCTCAGATGAAGTCAAGTGAAGT 409
QY 378 TGCTAAAGATCTCAAAAAGAGATGATGCAACCAAGAGTCTGACTGAGGCGAGTAGAGT 437
Db 410 AGCCAAACGCGAGTGAGCGTAGGCGACAGGAGAGTTCCTTGAAGCTGCCGAGGAGGC 469
QY 438 AGAGATAAAGATGCGATTGCAAAATTTGACGAGAGGAGCTGTAAAGGTCACAGGCAACA 497
Db 470 TGGAGAAGTGGAGAACAATTTGAGAAGTTCAATAAAGGCTTGTCAAGGTCAACCAAGCA 529
QY 498 CAACGAAGATTTGFAACGGCTATTAAGACTTATGGGGTTCTCTGTGTAGAGGCACTTC 557
Db 530 CATGAAGAGTGCAGAAATTCCTTAGCCTGATGGGTATTCCTTATGTGGATGCAACCTG 589
QY 558 TGAAGCAGAGCAGAAATGTGACGCCCTTTGCATAACGATAAGGTTTCGCTGTGCTTC 617
Db 590 TGAAGCTGAGGCCACATGCGCTGCCTTAGTAAAGCAGGAAAAGTGTATGCTGCGAGCAAC 649
QY 618 AGAAGATATGGACTCCCTTACHTTTTGGGCTCCAGGTTCTCTCGTCAATTAATGATCC 677
Db 650 TGAAGACATGGATGGCCCTGACCTTTGGCACTCCGCTGTCTAGGCACTCACAGCAAG 709
QY 678 AAGTTCCAAGAAATACCTGTGATGAAATTTGATGTGCAAGGTTTTGGAGGAGCTTGA 737
Db 710 TGAAGCTAAAGCTGCGCTATTTCAGAGTTCCTTAACCCGCTTTTTCAGGACATCGG 769
QY 738 ACTCACCATTGGACCAAGTTCATTTGATTGTGCACTCTGTGTGGATGTGACTATTTGTATG 797
Db 770 CATCAATCAGAACAAATTTGTGATCTCTGTATCTTGTGCGAAGTGAATTTATTTGTAAC 829
QY 798 CATCAAGGATGTCGGGGGCAACAGCTCTGAACCTTATTCGTCAACATGGGTCCATAGA 857

```

```

implicated in processing of Okazaki fragments during DNA
replication and in base excision repair; 5' nuclease"
/note="FEN-1 homolog; similar to mammalian Dnase IV or
MFL, to S. cerevisiae RAD27, RTH1 or YKL510, and to S.
pombe rad2; similar to Xenopus laevis xFEN1a encoded by
GenBank Accession Number U64563"
/codon_start=1
/product="xFEN1b"
/protein_id="AAB08478.1"
/db_xref="GI:1549393"
/translation="MGIHLAKLIADVAPAAIKHDIKSYFGRKVAVDASMCYIOFLI
AVRDGNMLQNEEGETTSHLMFKYIRLMEHGKIPVYVFDGPKQKSGSLAKRSE
RAAEAKLEEAEEAGEVENIEKFKRLVYKQHNECKKLLSLMGIPYVDPCEAE
ATCAALVKAGKYAAATDMDALTCTPVLRLHJASEAKLPIQEFHLNRYFDQIGI
NTEQFVDCILLGSDYCEITRIGIPKRAIDLJROHKTLEIINDIDNKKYPIPEWNLH
KEAROLFLEPEVIDADITELKWTEDDEGLVAFMGKEQFSEIRINRNGAKKLAKNRQG
STQGRLLDFFKVTGSSISSTKRKEVSGSKTKSKTKTGTPAGFKRKG"
BASE COUNT 461 a 288 c 377 g 343 t
ORIGIN

```

```

830 AATCCGTGGCATTGGACCCAAAGAGCCATTGACCTGATACGCGCAGATGAAGACTATTGA 889
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 858 AAGCATCTTGGAGAACTTAAATGAAGACAGATATCAAAATTCCTGGAGGACTGGCTTACCA 917
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 890 AGAGATCATGATAAATGACCTCAAGAAATACCCCATTCAGAGAACTGGCTGCACAA 949
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 918 AGAAGCTCGAGCGCTTGTTCAGGAGGCTAATGTCACA---TTGGATATTCCTGAGCTAAA 974
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 GGAGCCAGGACGCTTCTTGGAACACAGAGTGATAGATGCTGACATCACTGAGCTGAA 1009
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 975 ATGAGCTGACCTGATGAGGAGGCTCTATAAGTTTCTCGTGAAGAAATAAGTTTCAA 1034
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1010 GTGAGCTGAGCCTGATGAAGAGGAGCTGGTGGCTTCATGTGTGGGAGAGCAGTTTCA 1069
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1035 CGAAGATCGGCTGACAAAGCCATAGAGAGATCAAAATCTGCCAAGAAATAATTCGTCGA 1094
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 TGAAGACCGCATAGCAATGGGGCCAAAAGTTGGCCAAAGAACCGTCAAGGAGCAGCA 1129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1095 AGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACCA 1132
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1130 GGGCGGCTGGATGACTTTTCAAGTGAATGATCCA 1167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
XLU64563 1461 bp mRNA linear VRT 26-JUL-1999
LOCUS Xenopus laevis 5' nuclease xFEN1a mRNA, complete cds.
DEFINITION U64563
ACCESSION U64563
VERSION U64563.1 GI:1490869
KEYWORDS Xenopus laevis.
SOURCE Xenopus laevis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 1461)
Bibikova,M., Wu,B., Chi,E., Kim,K.H., Trautman,J.K. and Carroll,D.
Characterization of FEN-1 from Xenopus laevis. cDNA cloning and
role in DNA metabolism
J. Biol. Chem. 273 (51), 34222-34229 (1998)
99069415
PUBMED 9852084
2 (bases 1 to 1461)
Bibikova,M., Chi,E., Wu,B., Kim,K.-H. and Carroll,D.
Direct Submission
Submitted (19-JUL-1996) Biochemistry, U. of Utah Med. Sch., 50 N.
Medical Dr., Salt Lake City, UT 84132, USA.
Location/Qualifiers
1..1461
/organism="Xenopus laevis"
/db_xref="taxon:8355"
111..1259
/function="implicated in processing Okazaki fragments
during DNA replication and in base excision repair;
specifically removes single-stranded 5' sequence"
note="similar to FEN-1, DNaseIV, MFL1, RAD27, RTH1, and
rad2"
/codon_start=1
/product="5' nuclease xFEN1a"
/protein_id="AAB06176.1"
/db_xref="GI:1490870"
/translation="MGHGLAKLIADVAPAIKEHDIKSYFGRKRVAVDASMCYQFLI
AVROGNTLQNEEGEITSHLMGFYRTIRVHEGIRKPVYFDGKPPQMKSGELAKRSE
RRAEALKEAAEAGEVENIERKRLVKTQHNEEKKLLTLMGIPYVDAPCAE
ATCAALYKAGKVAAATEDMDALFTGTVLLRLHTASEAKKLPTQFHLNRVLDIGI
THEQFVLDLILGSDYCTETIRGIPKRAIDLIRHKTIDELIDNIDLKYPVFNHLH
KEAKHLEPEVDVTDITELKWIPEDEGLVPMCEKQFSEDRIRNGAKLKNROG
STQRLDDFPKVSSTTRKESGSAKKAKTKGGTPAGKPKRK"
454 a 293 c 365 g 349 t

BASE COUNT
ORIGIN
21.1%; Score 308.6; DB 5; Length 1461;
Query Match

```

```

Best Local Similarity 55.7%; Pred. No. 3e-66;
Matches 634; Conservative 0; Mismatches 499; Indels 6; Gaps 2;
.....QY 85 ATGGGCATCAAGGTTTACGAAACTGCTGCGGACATGCGCCCAAGCCGATGAAGAG 144
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 ATGGGAATTCAGGTTTGGCCAAACTTATCGCAGATGTGGCACCTGCGACGCTATCAAGAG 170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 145 CAGAAGTTTCGAGAGCTACTTCGGCCGCAAAATCGCCGTCGACGCCAGCATGAGCATATAC 204
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 CATGATATAAAAAGTTACTTTGGTCGAAAGTGCGGGTCGATGCTTCCATGTCATTAT 230
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 205 CAGTTCCTGATGTAGTTGGAGACAGCATGGAACCTCTCAAAATGAAGCTGGTGA 264
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 CAGTTTCTTATTCAGTCAGACAGGATGGCA---ACAGCTGCAAAAGAGAAAGTGA 287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 265 GTCAGTACTGATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTGGAAGCGGGAATC 324
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 ACCACAAGCCATCTTATGGGCATGTTTACCGTACCATTCGATGGTAGAGCATGGCATC 347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 325 AAGCCAGTTTATGTTTGTGGCAAGCCTCCTGATATGAAGAAACAAGAGCTTGTCTAAA 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 AAACCAAGTGTATGTTTGTGGAAAGCCACACAGATGAAGTCCGGTGAAGTAACTAGCTAAA 407
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 385 AGATACTCAAAAGAGATGATGCAACCAAAAGATCTGACTGAGGCAGTAGAGTAGAGAT 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 CGCAGTGAAGCTAGGCGGAGGACAGAGATTGCTTGAAGCTGCCGAGGAGGCTSGAGAA 467
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 AAAGATGCGATTGAAAATTTGAGCAAGAGGACTGTAAGGTCACAAGGCAACACACAGCAA 504
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 GTGGAGAACAATTGAGAAAGTTTCACTAAAGACTTGTCAAGGTTTCAAGCAGCACAATGAG 527
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 505 GATTGTAACCGCTTAAAGACTTATGGGGTTCCTGTTGTAGAGGACCTTCTGAAGCA 584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 GAGTGCAGAAATAATTACTTACCTAATGGGATTCCTTATGTGGATGACCTGTGAAGCT 587
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 565 GAAGCAAGATCTGAGCCCTTTTGATATAACGATAAGGTTTCCGCTGTGCTTCAGAGAT 624
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 GAGCCACGCTGCTGCTCTAGTAAAGCAGGAGGAAAGTGTATGCTGCAGCAACACAGAGAC 647
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 625 ATGAGCTCCCTTACTTTGGGGTCCACGGTTCCTTCGTCATTATATGATGCCAAGTTC 684
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 648 ATGGATGCCCTGACCTTTGGCACCCTGTCTGCTTAGGCACCTCAAGCA 707
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 685 AAGAAATACCTGTGAGGAATTTGATGTTGCCAAGTTTTGGAGGAGCTTGAAGTCAAC 744
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 708 AAAAGCTGCTATTCAGGAGTCCATCTAAACGCTGATTCAGGACATAGATATACACA 767
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 745 ATGACCAAGTTCATGATTTGTGCATCCTGTGGATGTGACTATTGTGATAGATACAAA 804
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 768 CATGAACAATTTGTGGATCTCTGTATCTTCTCGGAAGTGACTACTGTGAACCATCCGT 827
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 805 GGTATCGGGGGCAACAGCTCTGAAGACTTATTCGTCACATGGGTCCTATGAAGACATC 864
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 828 GGCATTCGACCCAAAAGAGCCATTGACCTGATCCGTCAGCATGAAGACATGACGAGATC 887
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 865 TTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAGGAGCTGGCTTACCAAGAGCT 924
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 888 ATTGATAACATTGACCTCAAGAAATACCCCGTCCCGGAAAACHTGGCTGCACAGGAGGCA 947
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 925 CGAGCGTTGTTCAAGGAGCCATAA---TGTCACATTTGGATATTCTTGAGCTTAAATGGACT 981
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 948 AAACATCTTTTCTTGGAAACCAAGTAGTCGACACTGACATCCAGGAGCTGAAGTGGATC 1007
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 982 GCACCTGATCAGAGGCTCTCATAGTTTCTCGTAAAGATAATGGTTTCAACGAGAT 1041
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1008 GAGCCGATGAAGAGGATTGGTGGCTTCATGTGTGGGGAGAAAGCAGTTTATG7GAAGAT 1067
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1042 CGGGTGACAAGGCCATAGAGAAATCAAAATCTGCCAAGAAATAAATCGTCGCAAGGAAGA 1101
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1068 CGCATACGCAACGGGCCAAAAGTTGGCCAAAGAACCCGCAAGGACACACAGGCGCGG 1127
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1102 CTCGAGTCTTTTTCAGGCCAAGTGCACACATAGCACCGCTAAACGGAAGGAGACT 1161
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 1128 CTGGATGACTTTTCAAAGTACCGGATCGTTAGCTCAACTAAAGAAAAGAGCGAGAA 1187
QY 1162 TCGGATAAACACGACGACCTCGACACAGAAACAAAGGCTGGTGGAAAGAGAA 1220
Db 1188 TCCAAAGGATCCGCTAAAGAAAGACCAAACTGGAGGTACCCCTGCAGGAAGTTCAA 1246

RESULT 8
AF065397
LOCUS
DEFINITION
Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.
ACCESSION
AF065397
VERSION
AF065397.1 GI:4106357
KEYWORDS
Xenopus laevis.
SOURCE
Xenopus laevis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1357)
Li, J.-L. and Cox, L.S.
Cloning and investigation of Xenopus FEN1: developmental expression
and function in DNA replication
Unpublished
JOURNAL
2 (bases 1 to 1357)
Li, J.-L. and Cox, L.S.
Direct Submission
AUTHORS
Submitted (14-MAY-1998) Department of Biochemistry, University of
Oxford, South Parks Rd., Oxford OX1 3QU, England
JOURNAL
Location/Qualifiers
FEATURES
source
1..1357
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1..1357
/gene="FEN1"
99..1247
/note="Involved in DNA replication and repair"
/codon_start=1
/product="flap endonuclease 1"
/protein_id="AAD02814.1"
/db_xref="GI:4106358"
/translation="MGHGLAKLIADVAPAIKEHDIKSYFGRKVVADSMCIYQFLI
AVRQDNTLQNEGETTSHLMGMFYRTIRWHEGKIPYVDFGRPPQMSSELAARSE
RAAEAKLEAEAEAGEVENIEKFKLVKVTYKOHNECKKLLITLGMIPYVDAPCEAE
ATCAALVRAKGVAAATDMATFTGTPVLLHLHTASEAKLPLOEFLNVRVQDIGI
TEOFVLCILLGSDYCEITIRGIGPKRALDIROKHTIDEIINDLKKYVPVENLH
KEAKHLFLEPVVDITELKWIPEDEGLVAFMGKOFSEDRIRNGAKKLAKNRQG
STQGRLLDFFKVTGVSYSSTKKEAESKSAKKKAKTGTTPAGKFKRGR"
BASE COUNT 425 a 284 c 345 g 303 t
ORIGIN
Query Match 21.0%; Score 307; DB 5; Length 1357;
Best Local Similarity 55.6%; Pred. No. 7.6e-66;
Matches 633; Conservative 0; Mismatches 500; Indels 6; Gaps 2;
QY 85 ATGGGCATCAAGGTTTTCAGAACTGCTGGCGGACAAATCGCCGCAAGGCGATGAGAG 144
Db 99 ATGGGAATTCACGGTTTGGCCAACTATTCGAGATGTGGCACCTGCAGCTATCAAGAG 158
QY 145 CAGAAGTTCGAGACTACTTTCGGCGCAAAATCGCGTCGACCGCATGAGCATATAC 204
Db 159 CATGATATAAAAGTTTACTTTGGTCGAAGAGTGGCGGTGCGATGCTTCATGTGCATTTAT 218
QY 205 CAGTTCTCTGATTGTAGTTGGAAGACAGAGGATGGAACCTCTCAAAATGAAGTGGTGA 264
Db 219 CAGTTTCTTATTCAGTGAGACAGAGATGGCA---ACACGTGCAAAACGAGGAAGTGA 275
QY 265 GTCACATGATTTGCAAGGATGTTCAACCGGCACATAGATCTACTGGAACGGGATC 324
Db 276 ACCACAAGCCATCTTATGGGCATGTTTACCCTGACCATTCGCATGGTAGGATGGCATC 335
QY 325 AAGCCAGTTTATGTTTTTGTGGCAAGCCCTCTGATATGAAGAAACAGAGCTTGTGTA 384
```

```
Db 336 AAACCAAGTGTATGTGTTGATGGAAAGCCACACAGATGAAGTCCGCTGAATAGCCAAA 395
QY 385 AGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGCGAGTAGAGGTAGAGAT 444
Db 396 CGAGTGAAGCTAGGCGGAGGAGAGAGTTCCTTGAAGCTGCCGAGGAGGTGGAGAA 455
QY 445 AAAGATGCGATTGAAAATTTGACAAAGAGAGACTGTAAAGGTCAACAGGCAACACAGAA 504
Db 456 GTGAGAAACATTGAGAAGTTTCACTAAAGACTTGTCAAGGTTACCAAGCAGCACAATAG 515
QY 505 GATTGTAAACGGCTATTAAACATTATGGGGTTCCTTGTGTAGAGGCACCTTCTGAAGCA 564
Db 516 GAATGCAAGAAATTTACTTACCCTTAATGGGCATTCCCTTATGTGGTGCACCTCTGAAGCT 575
QY 565 GAAGCAGAATGTGAGCGCTTTGCATAAAGATGAAGGTTCCTGCTTGTGCTTCAAGAGAT 624
Db 576 GAGCCACAGTGTGCTGCTAGTAAAGCAGGAGAAAGTGTATCTGCAGCAACAGAAGAC 635
QY 625 ATGGACTCCCTTACTTTTGGGGTCCACGGTTCCTCTGCTGCTCATTTAATGGATCCAAAGTTC 684
Db 636 ATGGATGCCCTGACCTTTGGCACCCCTGTCTCTAGGCACCTCACAGCAAGCGAAGCT 695
QY 685 AAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGAACCTCAC 744
Db 696 AAAAAAGCTGCTATTCAGGAGTTCCATCTAAACCGTGTCAATTCAGGACATAGTATCACA 755
QY 745 ATGACCAAGTTCATTGATTGTGTCATCCTGTGTGGATGTGACTATTGTATGATACATCAA 804
Db 756 CATCAACAATTTGGGATCTCTGTATCTTCTCGGAAGTGACTACTGTGAAGAACATCCGT 815
QY 805 GGTATCGGGGGGAAACAGCTCTGAAACTTATTTCGTCACATGGTCCATAGAAGCATC 864
Db 816 GGCATTGGACCAAAAGAGCAATTGACCTGATCGTCAGCATTAAGCATTCAGGAGATC 875
QY 865 TTGGAGAATCTTAATAAAGCAGATATCAAAATTCCTGAGGACTGGCTTACCAGAGACT 924
Db 876 ATTGATAACATTGACCTCAAGAAATACCCCTTCCGAAAACCTGGCTGCACAGGAGGCA 935
QY 925 CGACGCTTTGTTCAAGGAGCCCTAA---TGTCACATTTGATATTCTGAGCTTAAATGACT 981
Db 936 AAACATCTTTTCTTGGAACCAAGAGTAGTCACACTGACATCCAGCGAGCTGAAGTGGATC 995
QY 982 GCACCTGATGAGGAGGTCTCATAAGTTTCTGCTGTAAGATATATGTTTCAACGAAGAT 1041
Db 996 GAGCCCGATGAGAGAGATTGGTGGCCTTCATGTGTGGGAGAGAGCTTCAGTGAAGAC 1055
QY 1042 CGGFTGCAAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCCAGAGGAAGA 1101
Db 1056 CGCATACGCAACGGGGGCCAAAAGCTTGGCCAAAGAACCGGCAAGCAGCACACAGGGCGG 1115
QY 1102 CTCGAGTCTTTTCAAGCCCACTGCCACCATCATCAGCCGCTAAAACGGAAGGAGACT 1161
Db 1116 CTGGATGACTTTTTCAAAAGTACCGGATCCGTTAGCTCACTAAAGAAAAGAGGCGAGAA 1175
QY 1162 TCGGATAAAACAAAGCAGGAGCTGCGAACAAGAAAACAAAGGCTGGTGGAAAGAGAA 1220
Db 1176 TCCAAGATCCGCTTAAAGAAAGCCAAAACCTGGAGGTACCCCTGCAGGAAAGTTCAA 1234

RESULT 9
AF036327
LOCUS
DEFINITION
Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.
ACCESSION
AF036327
VERSION
AF036327.1 GI:2674206
KEYWORDS
Xenopus laevis.
SOURCE
Xenopus laevis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1149)
```

AUTHORS Kim,K., Biade,S. and Matsumoto,Y.
TITLE Involvement of flap endonuclease 1 in base excision DNA repair
J. Biol. Chem. 273 (15), 8842-8848 (1998)
MEDLINE 98204872
PUBMED 9535864
REFERENCE 2 (bases 1 to 1149)
AUTHORS Kim,K. and Matsumoto,Y.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) Radiation Oncology, Fox Chase Cancer
Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA
FEATURES
Location/Qualifiers
1..1149
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="oocyte"
1..1149
/gene="FEN1"
1..1149
/gene="FEN1"
/function="structure-specific nuclease"
/codon_start=1
/product="flap endonuclease 1"
/protein_id="AAB88707.1"
/db_xref="GI:2674207"
/translation="MGHGLAKLIADVAPAAIKKHEIDIKSYFGRKVAIDASMCYQFLI
AVRDGNFLONEGETTSHLMCMFYRTIRMVEHGKPVVFDGKPKOMKSGELAKRSE
RRAEAKLEAAEEAGEVENIEKFKRLVKVTKQHNBECKLLTILMGIPYVDPCEAE
ATCAALVAKAGKYAAATEMDALTFTGTVLLRLHTASEAKKLPIQEFHLNRVYQDIGI
THEQVDFCLILGSDYCEITIRIGIKPRAIDLRQHKTDIEIDNIDLKYPYVFNWLH
KEAKHLFDEPVVDITDELKVEPDEGLVAFMGCEKOFSEDRIRNKAFLAKNRQG
STQGRLLDPEKFTVGSVSTKRKEAGSKSAKKKAKTKGTGTPAGKFRKG"
BASE COUNT 352 a 247 c 306 g 244 t
ORIGIN
Query Match 20.9%; Score 305.4; DB 5; Length 1149;
Best Local Similarity 55.5%; Pred. No. 1.8e-51;
Matches 632; Conservative 0; Mismatches 501; Indels 6; Gaps 2;
QY 85 ATGGGATCAAGGGTTGACGAACCTCTCGCGGACAAATCGCCGACAGGCGATGAAGGAG 144
DB 1 ATGGGAATTCACGGTTTGGCCAACTTATCGCAGATGTGCACCTCGAGCTATCAAGAG 60
QY 145 CAGAAGTTTCAGAGCTACTTCTCGGCGCAAAATCGCGTCGACGCCAGCATGAGCATATAC 204
DB 61 CATGATATAAAGTTACTTTGTCGAAGAGTGGCGATCGATGCTTCCATGTGCAATTTAT 120
QY 205 CAGTTCCTGATGTAGTTGGAAGACAGGATGGAAGACTCTCAAAATGAAGCTGGTAA 264
DB 121 CAGTTTCTTATTTGGGTGAGACAGAGTGGCA---ACACGCTGCAAAACGAAGAGGTAA 177
QY 265 GTCACTAGTCAATTCGAAGGAATGTTCAACCGGACAAATAGATTACTGGAAGCGGAATC 324
DB 178 ACCAAGCCATCTATGGGTATGTTTACCGTACCAATCGCATGTTAGCAGCATGGCATC 237
QY 325 AAGCCAGTTTATGTTTGTGTCGAAGCTCTCTGATATGAGAACAAGAGCTTGTCTAAA 384
DB 238 AACACGATGATGTTTGTGGAAGCGCCGCCAGATGAAGTCGGGTGAACAGTACGCAAA 297
QY 385 AGATACTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGGTAGGAT 444
DB 298 CGCAGTGAGCGTGAAGGCGGACGAGAGAGTTGCTTGAAGCTGCGGAGGCGTGGAGAA 357
QY 445 AAGATGCGATTTGAAATTCAGCAAGAGGACTGTAAAGTGCACAAAGCAACACAAACAA 504
DB 358 GTGGAGAACATTTGAGAGTTCCTTAAAGAGCTTGTCAAGGTTTCCAGCACCAATGAG 417
QY 505 GATTGTAACCGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGACCTTCTTGAAGCA 564
DB 418 GAGTGCAGAAATTTACTTACCTTAATGGCATTCCTTATGTTGATGACACCGTGTGAAGCT 477
QY 565 GAAGCAGAAATGTGACGCCCTTTGCATAAACCAGTAAGGTGTTGCTGTTTCAGAAAT 624
DB 478 GAGGCCAGTGTGCTGCTTAGTAAAGCAGGGAAGTGTATGCTGCAGCAACAGAGAC 537

QY 625 ATGGACTCCCTTACTTTTGGGCTCCACGCTTCCCTTCTGTCATTATATGAT"CCAAGTTCC 684
DB 538 ATGGATGCGCTTACCTTGGCGCCCTGTGCTCCTTAGCCTCTACACCAAGCGAGCT 597
QY 685 AAGAAAAATACCTGTGTATGGAATTTGATGTTGCAAGGTTTGGAGGAGCTTGAAC"TCACC 744
DB 598 AAAAGCTGCGCTATTTCAGGAGTTCCATCTAAACCGGTGTCATTTCAGGACATAGGTATCACA 657
QY 745 ATGGACCACTTCAATGATTTGTCATCCTGTGTTGGATGTGACTATTTGATAGTACATCAAA 804
DB 658 CATGAACAATTTGGGATCTCTGTATCTTGTGCGGAAGTGACTACTGTGAACCATCCGT 717
QY 805 GGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCACATCGGTCCATAGAAAGCATC 864
DB 718 GGCATTGGACCCAAAAGAGCCATTGACCTGATCCGTCAGCATAAGACCATTCACGAGATC 777
QY 865 TTGGAGAATCTTAATAAGACAGATATCAAAATCTCTGAGGACTGGCTTACCAGAGCT 924
DB 778 ATTGATAACATTGACCTCAAGAAATACCCCGTTCCCGGAAACTGGGTGTCACAGGAGGCA 837
QY 925 CGAGCGCTTGTCAAGGAGGCTAA---TGTACATTTGGATATTCCTGAGCTTAAATGGACT 981
DB 838 AACATCTTTCTTGTGAACACAGAGTAGTGCACATCGACATCAGCAGCTGAAGTGGATC 897
QY 982 GCACCTGATGAGGAGGCTCTCATAGTTTCTGTTAAAAGATAATGTTTCAACAGAGAT 1041
DB 898 GAGCCCGATGAAGAAGGATTGGTGGCTTCATGTGTGGGAGAAGCAGTTTGTAGTGAAGAT 957
QY 1042 CGGGTGACAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTGCAGGAGGA 1101
DB 958 CGCATACGCAACGGGGGCAAAAGATTGGGCAAGAACCGGCAAGGCACACAGGCGCGG 1017
QY 1102 CTGAGTCTTTTTCAGCCCACTGCCACCATCATCGACCGCTTAAACGGAAGGAGACT 1161
DB 1018 CTGATGACTTTTCAAGTGACCGGATCGTTAGCTCACTAAAGAAAGAGGACAGAA 1077
QY 1162 TCGGATAAACAAGCAGGAGCTGCGAACAAGAAACAAGGCTGTTGGAAAGAGAA 1220
DB 1078 TCCAAAGGATCCCGTAAAGAAAGCCAAACTGGAGGTACCCCTGCGAAGAACTTCAA 1136
RESULT 10
BC010203 2163 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, Similar to flag structure-specific endonuclease,
DEFINITION clone MGC:7137 IMAGE:3158176, mRNA, complete cds.
ACCESSION BC010203
VERSION BC010203.1 GI:16307327
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 2163)
Strausberg,R.
Direct Submission
Submitted (05-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 36 Row: g Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers
 1. 2102

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="C57BL/6J"

/clone="MGC:28169 IMAGE:3985840"

/tissue_type="mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue."

/clone_lib="NCI-CGAP_Mam5"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

140. 1282

/codon_start=1

/product="unknown (protein for MGC:28169)"

/protein_id="AAH27295.1"

/db_xref="GI:20071071"

/translation="MGHGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYOFLL
 AVRGQVQNEGETTSHLMGFIRRMENGVRPVIVFDKPPQLKSGELARSE
 RRAEKLQQAQAGMEVEEVEKFLKLVKTRKOHNDCKHLLSLGIPYLDPEAE
 ASCALAKAGKYAAATDMDCTFGSPVLMRLHTASEAKLPIDFHLRSLVLELGL
 NQOFVDCILGSDYCESIRGIPRAVDLIQKHSEIEIIVRRDPSKYPVLENLH
 KEAQDLFEPLVDPESEVLEKWSPEEELVKPMCEKOFSEIRINSKVRILSKSRQ
 STQGRDLDFPVKVTGSLSSAKRKEPEKPGAPKAKTKAGGAKFRKK"

BASE COUNT 536 a 502 c 600 g 464 t

ORIGIN

Query Match 20.8%; Score 303.8; DB 10; Length 2102;

Best Local Similarity 56.2%; Pred. No. 5.2e-65;

Matches 641; Conservative 0; Mismatches 482; Indels 18; Gaps 3;

85 ATGGGCAATCAAGGGTTGACAACTGCTGGCGGACATGCGCCCAAGCGCATGAAGGAG 144

Db 140 ATGGGAATTCACGGCTTCCCAACTAATGCTGATGTGGCCCGCCAGTGGCATCCGTGAG 199

QY 145 CAGAAGTTCAGAGCTACTTTCGGCGCAAAATCGCGCTGACGCGCAGCATGAGCATATAC 204

Db 200 AATGACATCAAGAGCTACTTTCGGCGCAAAATCGCGCTGACGCGCAGCATGAGCATATAC 259

QY 205 CAGTCTCTGATGTAGTGGAGGACAGCATGGAACCTCTCAACATGAAGCTGGTGA 264

Db 260 CAGTCTCTGATGTAGTGGAGGACAGCATGGAACCTCTCAACATGAAGCTGGTGA 316

QY 265 GTCACTAGTCAATTCGAAGGAATGTTCAACGGCAATAAGATTACTGGAAGCGGGAATC 324

Db 317 ACCACGACCACTGATGGCATGTTCTACCGTACCCTCCGATGATGGAGATGGCGTC 376

QY 325 AAGCCAGTTATGTTTGTATGCAAGCTCTCTGATATGAAGAACAGAGCTTGTCTAAA 384

Db 377 AAGCTGTGTACCTCTTTGATGCAACACCAACAGCTGAGTCAAGGCGCTGGCCCAAG 436

QY 385 AGATACTCAAAAAGATGATGACCAACCAAGATCTGACTGAGCGCATAGAGGTAGGAGAT 444

Db 437 CGCAGTGAGAGCGCGCCGCGGCTGAGAGCAACTGACGAGCGCTCAGGAGCGCTGGGATG 496

QY 445 AAGATCGGATCAAAAATTTAGCAGAGGAGCTGTAAGGCTCAAGGCAACACACACGAA 504

Db 497 GAGGAGAGTGGAGAGTTCACCAAGAGGCTGCTGAGAGGTCAACACACACACATGAT 556

QY 505 GATTGTAACGGGTATTAAGACTTATFGGGGGTTCCTGTGTAGAGGCACTTCTGAAGCA 564

Db 557 GAGTGCAAAACACCTGTGTAGCGCTCATGCGGCATCCCTTACCTTGTATGCAACCGAGGCA 616

QY 565 GAAGCAAGATGTCAGCCCTTTTCATATAAAGATAGGTTGCTGCTGTGTTCTGAGAGAT 624

Db 617 GAGCCAGCTGTGCTGCCCTGGCAAGGCTGGCAAGCTATGCTGCGGCCACCGAGGAC 676

QY 625 ATGGACTCCCTTACTTTTGGGGCTCCACCGTTCCTTCCTTCATTTAATGGATCAAGTTC 684

Db 677 ATGGACTCGCTCACTTTTGGCAGCCCGCTGCTAATGCGACACTTAACCTGCCAGTGAGGCC 736

QY 685 AAGAAATACCTCTGATGGAATTTGATGTTGCCAAGGTTTTCGAGGAGCTTGAACCTACCC 744

Db 737 AAGAAAGTGCCTCAACCAAGAGTTCCATCTGAGCGCGCTCCTCAGAGAGCTGGGTCTGAAC 796

QY 745 ATGACCAAGTTCATGATTTGTCATCCCTGTCGTGATGAGTGAAGTATTTGATAGATCAAA 804

Db 797 CAGGAGCAGTTTGTGGATCTGTGCATCTCTGCTGGGTAGCGACTACTTGGGAGAGCCTCGT 856

QY 805 GGTATCGGGGGGCAACACAGCTCTGAACCTTATTGCTCAACATGCGTCCATAGAAAGCATC 864

Db 857 GGCATTGGGCCCAAGCGGCTGTGGATCTCATCCAGAAACATAAGAGCATCAGGAGATC 916

QY 865 TTGGAGAAATCTTAAATAAGACAGATATCAATTCCTGAGGAGCTGGCCTTACCAAGAACT 924

Db 917 GTGAGCGGCTGGACCCAGCAAGTACCCCGTTTCCAGAGAACTGGCTCCACAAGGAAGCC 976

QY 925 CGAGCGCTGTTCAGAGGAGCCTAATG--TCACATTTGGATATTCCTGAGCTAAATGAGACT 981

Db 977 CAGCAGCTCTTCTTGGAGCCAGAACTGAGCCAGAGCTCTGTGAGCTGAAGTGGAGC 1036

QY 982 GCACCTGATGAGGAGGCTCTCATAGCTTCCCTGTTAAAGATAATGTTTCAACGAAGAT 1041

Db 1037 GAGCAAAATGAAGAGAGTTGGTCAAAATTTGTTGGTGAAGAGCAGTTTCTGAAGAG 1096

QY 1042 CGGTGCAAAAGGCCATAGAGAAGATCAAACTGTCCCAAGAAATAAATCGTCGCAAGGAAGA 1101

Db 1097 CGAATTCGCAAGTGGGTCAAGCGGCTGAGTAAGAGCGCCAGGCGAGCACCACGAGGAGC 1156

QY 1102 CTCGAGTCTCTTTTCAAGCCAACTGCGCACATCAGCACCGCTAAACGGAAGGAGACT 1161

Db 1157 CTCGATGATTTCTTCAGGTTGACAGGCTCAGCTCTCTCTCAGC-----TAAGCGC 1204

QY 1162 TCGATTAACCAAGCAGGAGCTGGACACAGAAACAAAGCTGCTGGTGAAGAGAA 1221

Db 1205 AAGAGCGCAACCAAGGGGCTGCTAAGAGAAAGCAAGACTGGGGGCGGGGAG 1264

QY 1222 T 1222

Db 1265 T 1265

RESULT 13

HSRAD2 Homo sapiens mRNA for flap endonuclease-1. 1757 bp mRNA linear PRI 07-DEC-1998

LOCUS X76771

DEFINITION X76771.1 GI:1905802

ACCESSION X76771.1

VERSION 2

KEYWORDS fln1 gene; flap endonuclease-1.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Watts, F.

1 (bases 1 to 1757)

Direct Submission

Submitted (16-DEC-1993) F. Watts, University of Sussex, School of

Biological Sciences, Falmer, Brighton BN1 9QJ, UK

2 (bases 1 to 1757)

Murray, J.M., Tavassoli, M., al-Harithy, R., Sheldrick, K.S.,

Lehmann, A.R., Carr, A.M. and Watts, F.Z.

Structural and functional conservation of the human homolog of the

Schizosaccharomyces pombe rad2 gene, which is required for

chromosome segregation and recovery from DNA damage

Mol. Cell. Biol. 14 (7), 4878-4888 (1994)

94277093

PUBMED 8007985

Location/Qualifiers

1. 1757

/organism="Homo sapiens"

/db_xref="taxon:9606"

source

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 1 Row: 9 Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 1905802.

FEATURES

source

1..2031
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="LocusID:2237"
 /db_xref="taxon:9606"
 /clone="MGC:8478 IMAGE:2821792"
 /tissue="lung, small cell carcinoma"
 /clone_lib="NIH_MGC_7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 139..1281

CDS

/codon_start=1
 /product="flap structure-specific endonuclease 1"
 /protein_id="AAH00323.1"
 /db_xref="GI:12653113"
 /translation="MGIGLAKLIADVAPSAIRENDIKSVFGRKVAIDASMSIVQFLI
 AVROGDVLDNEGETTSHLMFYRTIRMMENIKPVYFDGKPPOLKSGELAKRSE
 RRAEAKLOQAAGABOEVEKFTKRLVYKOHNDCKHLPLQEFHLSLPIYLDAPSAAE
 ASCAALVKAGYAAATDMDCLEFGSPVLMHLLTASEAKLPIQEFHLSLPIYLDAPSAAE
 NOEYVDLCILLSDYCSIRGIGPKRAVDLIQKHSIEBIVBELDNPKNKYVPELNHL
 XEAHLFLEPDEPVSVELKWSPEEELIKFMCEKQFSEERIRSGVRLSKSRG
 STQGRDFFKVTGSLSSAKRKEPEPKGTSKKAKTGAAGKFRKG"
 533 a 465 c 589 g 444 t

BASE COUNT

ORIGIN

Query Match 20.7%; Score 302.8; DB 9; Length 2031;

Best Local Similarity 55.0%; Pred. No. 9, 2e-65;

Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;

QY 85 ATGGGATCAAGGGTTGACGAACTGTCGGCGACATGCGCCCAAGCCGATGAGGAG 144
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 139 ATGGGAATTCAGGCTGGCCAACTAAATGCTGTGTCGGCCCGCCAGTGCCTCGGGAG 198
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 145 CAGAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCGCTCGACGCCAGCATGACATATAC 204
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 199 AATGACATCAAGAGCTACTTTGGCCGTAAGGTGGCCATTGATGCTCTATGACCATTTAT 258
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 205 CAGTTCTGATCTAGTCTGGAAGGACGCGATGCAAACTCTCACAAATGAAGCTGGTAA 264
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 259 CAGTTCTGATCTGCTGCGCAG---GGTGGGATGCTGCAAGATGAGGAGGCTGAG 315
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 265 GTCAGTATCATTTGCAAGGAATGTTCAACCGGACAAATTAAGATTACTGGAAGCGGAATC 324
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 316 ACCACGAGCCAGCTGATGGCATGTTCTACCGCACCATTCGCATGATGAGAACGCGCATC 375
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 325 AAGCCAGTTTATGTTTGTATGGCAAGCTCTGATGATGAAGAAACAAGAGCTTCTAAA 384
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 376 AAGCCGCTGATGCTTTGATGCGAAGCCGCCACAGCTCAAGTCAGGCGAGCTGGCCAA 435
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 385 AGATCTCAAAAAGATGATGCAACCAAGATCTGACTGAGGCGAGTAGAGTAGAGAT 444
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 436 CGCAGTGAAGCGGCGCTGAGCGAGAGAGAGCTGACAGAGCTCAGCTCTGGGCC 495
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 445 AAGATCGGATTTGAAAATTTGACAGAGAGAGCTGTAAAGGTCACAAAGGCAACACACGAA 504
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 496 GAGCAGAGGTGAAAATTTCAAGAGCGCTGTTGAAGGTCACTAAGCAGCACAATGAT 555
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 505 GATTGTAACGGCTATTAAGACTTATGGGGTTCCTGTTGTAGAGGCACTTCTGAGCA 564
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 556 GAGTGAACATCTGCTGAGCCCTCATGGGATCCCTTATCTGATGACCCAGTAGGCA 615
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 565 GAAGCAGAATGTCAGCCCTTTGCATAAAGATGATGAGGTTTCGCTGTTCTTCAAGAAT 624
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 616 GAGCCAGCTGCTGTCCTGGTGAAGGCTGGCAAAAGTCTATGCTCGGCTACCGAGGAC 675
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 625 ATGACCTCCCTTACTTCTGGGCTCCACGCTTCTCTGCTATTTAATGATCCAGTCC 684
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 676 ATGACCTGCCTTACCTTTCGGACGCCCTGTGCTAATCGACACCTGACTGCCAGTGAAGCC 735
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 685 AAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGAATCACC 744
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 736 AAAAGCTGCCATCCAGGAATTCACCTGAGCCGGATTCTCAGGAGCTGGGCTGAAC 795
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 745 ATGAGCAGCTTCAATGATTTGTCATCCTGTCGATGTCGATGTCGATGTCGATGTCGAT 804
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 796 CAGGAACAGTTTGGATCTGTCATCTCTAGGAGTCTCTGAGGAGTCTCTGAGGAGTCTCGG 855
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 805 GGTATCGGGGGGCAACAGCTCTGAACATTTATCGTCAACATGGGTGCCATGAGAAGCATC 864
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 856 GGTATCGGGGGGCAACAGCTCTGAACATTTATCGTCAACATGGGTGCCATGAGAAGCATC 915
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 865 TTGAGAGATCTTAAATAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCT 924
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 916 GTGCGGAGCTTGACCCCAACAGTACCTGTGCCAGAAATTTGGCTCCACAAGAGGCT 975
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 925 CGACGCTGTTCAAGGAGCTTATG---TCATGTTGATATTCCTGAGCTAAATGGACT 981
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 976 CACGAGCTCTTCTTGAACCTGAGTGTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 1035
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 982 GCACCTGATGAGGAGGCTCTCATTAAGTTTCTGTTAAAGATAAATGGTTTCAACGAAGAT 1041
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1036 GAGCAAAATGAAGAGAGCTGATCAAGTTTCAATGTGTGGTGAAGAGCAGTTCTCTGAGGAG 1095
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1042 CGSGTGACAAAGCCATAGAGAGATCAAAATCTGCCAAGAAATAAATCGTCGCAAGGAGA 1101
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1096 CGAATCGCAGTGGGTCAAGAGCTGAGTAAGAGCCGCAAGGAGCAGCCAGGCGCGC 1155
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1102 CTCGAGTCTTTTCAAGCCAACTGCCAC---CACATGACACCGCTAAACGGAGGAG 1158
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1156 CTGATGATTTCTTCAAGGTGACCGCTCAGTCTCTTCAAGCGAAGGAGGACGAGAA 1215
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1159 ACTTCGATATAAAGAGGAGCTGCGAACAACAAAGGCTGCTGGAAGAG 1218
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1216 CCAAGGATCCACTAAGAGAGGCAAGAGCTGGGCGAGCGGAAGTTTAAAGGGGA 1275
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1219 AATAATCTTGAATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1278
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1276 AATAATGTTTTCCTCCATATACCTCTTCCACCCAGAAATTTTCCGCTTTTACCC 1335
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1279 TT 1280
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1336 TT 1337
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15

AF523117 AF523117 6536 bp DNA linear PRI 30-JUL-2002
 LOCUS Homo sapiens flap structure-specific endonuclease 1 (FEN1) gene,
 complete cds.
 DEFINITION
 AF523117
 ACCESSION
 AF523117 GI:21668122
 KEYWORDS
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 6536)
 AUTHORS
 Rieder, M.J., Livingston, R.J., Braun, A.C., Montoya, M.A.,
 Chung, M.-W., Miyamoto, K.E., Nguyen, C.P., Nguyen, D.A., Poel, C.L.,
 Robertson, P.D., Schackwitz, W.S., Sherwood, J.K., Witrak, L.A. and
 Nickerson, D.A.
 DIRECT SUBMISSION
 Submitted (20-JUN-2002) Genome Sciences, University of Washington,
 1705 NE Pacific, Seattle, WA 98195, USA
 COMMENT
 To cite this work please use: NIHES-SNPs, Environmental Genome
 Project, NIHES ES15478, Department of Genome Sciences, Seattle, WA
 (URL: <http://epg.gs.washington.edu>).
 FEATURES
 Location/Qualifiers
 source
 1..6536
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 variation 103

```
/frequency="0.01"  
/replace="c"  
360  
variation  
repeat_region  
/frequency="0.01"  
/replace="t"  
393.456  
/rpt_family="L2"  
/rpt_type-dispersed  
504  
variation  
/frequency="0.01"  
/replace="t"  
974  
variation  
/frequency="0.01"  
/replace="t"  
982  
variation  
/frequency="0.01"  
/replace="c"  
995  
variation  
/frequency="0.40"  
/replace="a"  
1064.5624  
gene  
/gene="FEN1"  
mRNA  
Join(1064.1414,3727.5624)  
/product="flap structure-specific endonuclease 1"  
1175  
variation  
/gene="FEN1"  
/frequency="0.07"  
/replace="t"  
1478.1628  
repeat_region  
/rpt_family="MIR"  
/rpt_type-dispersed  
1826.1971  
repeat_region  
/rpt_family="MIR"  
/rpt_type-dispersed  
2202  
variation  
/gene="FEN1"  
/frequency="0.02"  
/replace="g"  
2345.2510  
repeat_region  
/rpt_family="Alu"  
/rpt_type-dispersed  
2568.2787  
repeat_region  
/rpt_family="Alu"  
/rpt_type-dispersed  
2632  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="a"  
2907.3006  
repeat_region  
/rpt_family="MER1_type"  
/rpt_type-dispersed  
3138  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="g"  
3276.3319  
repeat_region  
/rpt_family="ERV1"  
/rpt_type-dispersed  
3690  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="g"  
3748.4890  
CDS  
/gene="FEN1"  
/codon_start=1  
/product="flap structure-specific endonuclease 1"  
/protein_id="AM74238.1"  
/db_xref="GI:2168423"  
/translation="MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIQFLI  
AVRGGVLDNDEGETSHLMGMFYETIRMENGIKPVYFDGKPPOLKSGELAKRSE  
RRAPAEKQQAQAQAQAEQVEKFTKRLVKVTQKHNDCEKHLLSLMGPIYLDAPSEAE  
ASCAALVKGKVAATAEDMDCLTFGSPVLMRHLTASEAKKLP1QEPHLSRLQELGL
```

```
misc_feature  
/gene="FEN1"  
/note="region not scanned for variation"  
4246.4405  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="g"  
4920  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="t"  
4973  
variation  
/gene="FEN1"  
/frequency="0.01"  
/replace="t"  
5213  
variation  
/gene="FEN1"  
/frequency="0.41"  
/replace="t"  
6108  
variation  
/frequency="0.03"  
/replace="t"  
6346  
variation  
/frequency="0.01"  
/replace="c"  
BASE COUNT 1701 a 1441 c 1706 g 1688 t  
ORIGIN  
Query Match 20.7%; Score 302.8; DB 9; Length 6536;  
Best local Similarity 55.0%; Pred. No. 1.1e-64;  
Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;  
...Qy 85 ATGGCATCAAGGGTTTGACGAACCTGCTGGCGGACAATGTCGCCCAAGGCGCATGAAGAG 144  
Db 3748 ATGGCAATCAAGGCTGCGCAAACTAATTGCTGATGTGGCCCCCAGTGCATCCGGGAG 3807  
Qy 145 CAGAACTCGAGAGCTACCTGGCGGCAAAATCGCGTCGCGCAGCAGCATGATATAC 204  
Db 3808 AATGACATCAAGAGCTACCTGGCGGCAAAATGCTGATGTGGCCCCCAGTGCATCCGGGAG 3807  
Qy 205 CAGTCTGATTTGTTGAAGGACAGCATGGAACCTCTCAAAATGAAGCTGGTGA 264  
Db 3868 CAGTCTGATTTGTTGCGGAG--GGTGGGATGTGCTGCAGAAATGAGAGGGTGAG 3924  
Qy 265 GTCACTAGTCAATTTGCAAGGAATGTTCAACGGGACAATAAGATTACTGGAAGCGGAATC 324  
Db 3925 ACCACAGCCACCTGATGGCGATGTTACCGCACCATCGCATGATGGAGACGGCATC 3984  
Qy 325 AAGCAGTTTATCTTTTATGGCAAGCCTCTGATATGAGAAACAGAGCTTGTCTAA 384  
Db 3985 AAGCCGCTGATGCTTTTATGGCAAGCCTCTGATATGAGAAACAGAGCTTGTCTAA 4044  
Qy 385 AGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGAGTAGAGTAGAGAT 444  
Db 4045 CGCAGTGAGCGGGGCTCAGCAGAGAAGCAGCTGCAGCAGGCTCAGGCTGCTGGGGCC 4104  
Qy 445 AAGATCGGATTCAAAATTTGAGCAAGAGGAGCTGTAAGGTCACAAGGCAACACAA 504  
Db 4105 GAGCAGAGGTGAAAATTTCACTAAGCGGCTGTGAGGTCACCTAAGCAGCACTGAT 4164  
Qy 505 GATTGTAAGCGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGACCTTCTGAAGCA 564  
Db 4165 GAGTGCAAAACATCTGCTGAGCCTCATGGCATCCCTTATCTTGTATGCACCCAGTGAGCA 4224  
Qy 565 GAAGCAGAAATGTGACGCCCTTTGCATAAACGATAGGTGCTGCTGCTCAGAGAT 624  
Db 4225 GAGCCAGCTGTGCTGCCCTGCTGAGGCTGGCAAGTCTATGCTGGGCTACCGAGGAC 4284  
Qy 625 ATGACTCCCTTACTTTTGGGCTCCACGGTTCCTTCCGTCATTTAATGGATCCCAAGTCC 684  
Db 4285 ATGACTGCTCCTACCTTGGGAGCCCTGTGTAATTCGACACCTGACTGCCAGTGAAGCC 4344  
Qy 685 AAGAAAATACCTGTGATGAAATTTGATGTGCAAGGTTTGGAGGAGCTTGAACCTCAC 744
```

```
Db 4345 AAAAGCTGCCAATCCAGGAATCCACCTGAGCCGGATTCTGCAGGAGCTGGGCCCTGAAC 4404
QY 745 ATGGACAGTTTCATTGATTGTGCATCCCTGTGTGGATGTGACTATTGTGATAGCATCAAA 804
Db 4405 CAGGAACAGTTTGTGGATCTGTGCATCCCTGTAGGCAGTGTACTACTGTGAGAGTATCCGG 4464
QY 805 GGTATCGGGGGGCAACACAGCTCTGAACCTTATTCGTCAACATGGGTCCATAGAAAGCATC 864
Db 4465 GGTATTGGGGCCCAAGCGGCTGTGGACCTCATTCCAGAACCAAGAGCATCGAGGAGATC 4524
QY 865 TTGGAGAACTTTAATAAGACAGATATCAAAATTCCTGAGGACTGGCCTTTACCAAGAAGCT 924
Db 4525 GTCGGGGACTTGACCCCAACAAGTACCCTGTGCCAGAAATTTGCTCCACAAGGAGCT 4584
QY 925 CGAGGCTTGTCAAGGAGCTATG---TCACATTGGATATTCTTGAGCTAAATGGACT 981
Db 4585 CACCAGCTCTTCTTGAACCTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 4644
QY 982 GCACCTGATGAGGAGGTCTCATAGTTTCCCTGGTAAAGATAATGGTTTCAACGAAGAT 1041
Db 4645 GAGCCAAATGAAGAAGAGCTGATCAAGTTTCATGTGTGTGAAAGCAGTTCTCTGAGGAG 4704
QY 1042 CGGGTGACAAGGCCATAGAGAATCAAAATCTGCCAAGAAATAATCTCGCAAGGAAGA 1101
Db 4705 CGAATCCGAGTGGGTCAAGAGGCTGAGTAAGAGCCGCCAAGGCGAGCACCAGGCGCGC 4764
QY 1102 CTCGAGTCTCTTTTCAAGCCAACTGCCAC---CACATCAGCACCGCTAAACCGGAAGGAG 1158
Db 4765 CTGGATGATTCTCAAGGTGACCGGCTCAGTCTCTTCAGCTAAGCGCAAGGAGCCAGAA 4824
QY 1159 ACTTCGGATAAACAAGCAGGAGCTGCGACAAGAAACAAAGGCTGGTGGAAAGAAG 1218
Db 4825 CCAAGGGATCCACTAAGAGAAGGCAAGAGCTGGGGCAGCAGGGAAGTTTAAAGGGGA 4884
QY 1219 AATAATCTTTGGATGTTGATGTACAACTACGAAAGCAGCGGTGGCGTGATCAC 1278
Db 4885 AATAAATGTGTTCCCATTTACCTCTTACCCAGNATATTTGCCGCTCTTGATCCC 4944
QY 1279 TT 1280
Db 4945 TT 4946
```

Search completed: May 8, 2003, 21:02:13
Job time : 3813 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 19:38:47 ; Search time 360 Seconds
(without alignments)
9151.871 Million cell updates/sec

Title: US-09-805-311-1
Perfect score: 1463
Sequence: 1 cagcagaatagctcgccgccc.....aaaaaaaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*			
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*		
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*		
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*		
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*		
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*		
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*		
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*		
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*		
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*		
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*		
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*		
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*		
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*		
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*		
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*		
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*		
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*		
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*		
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*		
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*		
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*		
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*		
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*		
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	21 AAA27923	Maize Rad2/FEN-1 c
2	1421.8	97.2	1478	21 AAA27926	Maize Rad2/FEN-1 c
3	1409.8	96.4	1541	21 AAA27924	Maize Rad2/FEN-1 c
4	1350.2	92.3	1381	21 AAA27925	Maize Rad2/FEN-1 c
5	302.6	20.7	1144	20 AAX02117	Human FEN-1 DNA fr
6	302.6	20.7	1144	20 AAX02107	Human FEN-1 cDNA
7	276.6	18.9	2033	20 AAX02111	Human FEN-1 genomi
8	276	18.9	1300	20 ABL14287	Drosophila melanog
9	272.2	18.6	1930	20 AAX02108	Mouse FEN-1 cDNA

10	270.2	18.5	1149	20 AAX02109	Yeast FEN-1 cDNA
11	252.6	17.3	3362	23 ABL14286	Drosophila melanog
12	173.2	11.8	5471	23 ABL10094	Drosophila melanog
13	159.6	10.9	1023	18 AAT76685	Pyrococcus furiosu
14	159.6	10.9	1023	19 AAV63840	Pyrococcus furiosu
15	159.6	10.9	1023	19 AAV53951	Nucleotide sequenc
16	152.8	10.4	1054	19 AAV53984	P. furiosus N-term
17	149.6	10.2	1115	19 AAV53992	P. furiosus N-term
18	139.6	9.5	510	22 AAK37479	Human bone marrow
19	139.6	9.5	510	22 AAL18286	Probe #8219 for ge
20	135.6	9.3	1032	20 AAX31850	Pyrococcus heat re
21	134	9.2	349980	22 AAH41224	Pyrococcus abyssi
22	128	8.7	514	19 AAV53990	Nucleotide sequenc
23	126.4	8.6	1164	19 AAV53975	Nucleotide sequenc
24	125.6	8.6	1115	19 AAV53978	P. furiosus N-term
25	117.6	8.0	1729	23 AAS83732	DNA encoding novel
26	110.2	7.5	889	19 AAV53974	Nucleotide sequenc
27	107.8	7.4	417	24 ABN77330	Human nuclease-lik
28	98.6	6.7	1011	19 AAV65913	Archaeoglobus fulg
29	98.6	6.7	1011	19 AAV54009	Nucleotide sequenc
30	97.6	6.7	1164	19 AAV53989	M. jannaschii N-te
31	94.8	6.5	386	19 AAV53979	Human foetal liver
32	90	6.2	546	22 ABA63065	Human brain expres
33	90	6.2	546	22 AAK11490	Human bone marrow
34	90	6.2	546	22 AAK37269	Human bone marrow
35	90	6.2	546	22 AAL18098	Probe #8031 for ge
36	90	6.2	546	24 ABS11263	Human genome-deriv
37	89.2	6.1	981	18 AAT76682	Methanococcus jann
38	89.2	6.1	981	19 AAV53948	Nucleotide sequenc
39	89.2	6.1	1664976	19 AAV21209	Methanococcus jann
40	86.6	5.9	245	22 AAK50354	Human bone marrow
41	86.6	5.9	245	22 AAI27416	Probe #17349 for g
42	86	5.9	981	19 AAV65837	Methanococcus jann
43	76.2	5.2	777	19 AAV54010	Nucleotide sequenc
44	74.6	5.1	840	19 AAV53977	Nucleotide sequenc
45	69.4	4.7	987	19 AAV54011	Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAA27923	standard; cDNA; 1463 BP.
ID	AAA27923
XX	AAA27923;
XX	AC
XX	12-SEP-2000 (first entry)
DT	Maize Rad2/FEN-1 cDNA.
DE	Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX	endonuclease; exonuclease; DNA repair; gene targeting; ss.
KW	Zea mays.
OS	
XX	Key
XX	Location/Qualifiers
FT	85..1224
FT	/*tag= a
XX	WO200036109-A1.
PN	
XX	22-JUN-2000.
PD	
XX	16-NOV-1999; 99WO-US27147.
XX	15-DEC-1998; 98US-0112332.
PR	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	MahaJan PB;
XX	WPI; 2000-452026/39.

DR P-PSDB; AAY95307.

XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
PS Claim 1; Page 69-71; 85pp; English.
XX

CC The present sequence is that of maize cDNA coding for RAD2/FEN-1
CC (see AAY95307). The corresponding RNA was isolated from immature
CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.

XX SQ Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;

Query Match 100.0%; Score 1463; DB 21; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CACGAGAATAGCTCGGGCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCGCCGCCA	60
DB	1	CACGAGANTAGCTCGGGCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCGCCGCCA	60
QY	61	CCGCGCACAGCCCGCCGACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGAC	120
DB	61	CCGCGCACAGCCCGCCGACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGAC	120
QY	121	AATGCGCCCAAGCGGATGAAGGACGAGATGTCGAGAGCTACTCGCGCGCAAAATCGCC	180
DB	121	AATGCGCCCAAGCGGATGAAGGACGAGATGTCGAGAGCTACTCGCGCGCAAAATCGCC	180
QY	181	GTGACGCCAGCATGACATATACAGTTCCTGATTTGATGGAGACAGGATGAA	240
DB	181	GTGACGCCAGCATGACATATACAGTTCCTGATTTGATGGAGACAGGATGAA	240
QY	241	ACTCTCACAAATGAAGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACA	300
DB	241	ACTCTCACAAATGAAGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACA	300
QY	301	ATAAGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCCTCCTGAT	360
DB	301	ATAAGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCCTCCTGAT	360
QY	361	ATCAAGAAACAGAGCTTCTTAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTG	420
DB	361	ATCAAGAAACAGAGCTTCTTAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTG	420
QY	421	ACTGAGGCAGTAGAGTAGGATGAAGATGCGATTGAAATTTGAGCAAGGAGCTGTA	480
DB	421	ACTGAGGCAGTAGAGTAGGATGAAGATGCGATTGAAATTTGAGCAAGGAGCTGTA	480
QY	481	AAGGTACAGAGGCAACACAGCAAGATTTGTAACGGCTATTAAAGACTTATGGGGTTCCCT	540
DB	481	AAGGTACAGAGGCAACACAGCAAGATTTGTAACGGCTATTAAAGACTTATGGGGTTCCCT	540
QY	541	GTGTGTAGAGGCCTTCTTCAAGCAGAGCAAGATGTCAGCCCTTTGCATAAACGATAAG	600
DB	541	GTGTGTAGAGGCCTTCTTCAAGCAGAGCAAGATGTCAGCCCTTTGCATAAACGATAAG	600

QY	601	GTGTTGCGTGTGCTTCAGAGATATGAGCTCCCTTACTTTTGGGGCTCCACAGGTTCCCTT	660
DB	601	GTGTTGCGTGTGCTTCAGAGATATGAGCTCCCTTACTTTTGGGGCTCCACAGGTTCCCTT	660
QY	661	CGTCATTATATGATCCCAAGTTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAG	720
DB	661	CGTCATTATATGATCCCAAGTTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAG	720
QY	721	GTGTTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTTGATTTGTGATCCTCTGTGGA	780
DB	721	GTGTTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTTGATTTGTGATCCTCTGTGGA	780
QY	781	TGTGACTATTTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAACATTTATTCGT	840
DB	781	TGTGACTATTTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAACATTTATTCGT	840
QY	841	CAACATGGTCCATAGAAGCATCTTGGAGATCTTTAATAAAGACAGATATCAAAATTCCT	900
DB	841	CAACATGGTCCATAGAAGCATCTTGGAGATCTTTAATAAAGACAGATATCAAAATTCCT	900
QY	901	GAGGACTGCCCTTCAACAGAGCTCGAGCTTGTTCAGGAGCCTTAATGTCACATTCGAT	960
DB	901	GAGGACTGCCCTTCAACAGAGCTCGAGCTTGTTCAGGAGCCTTAATGTCACATTCGAT	960
QY	961	ATTCTGTAGCTAAAATGGAGCTGCACCTGATGAGGAGGCTCTCATAAGTTTCTGTTAAAA	1020
DB	961	ATTCTGTAGCTAAAATGGAGCTGCACCTGATGAGGAGGCTCTCATAAGTTTCTGTTAAAA	1020
QY	1021	GATATGTTTCAACAGAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAG	1080
DB	1021	GATATGTTTCAACAGAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAG	1080
QY	1081	ATAAATCTGTCACAGGAGACTCGAGTCTTTTCAAGCCCACTGCCACCATCAGCA	1140
DB	1081	ATAAATCTGTCACAGGAGACTCGAGTCTTTTCAAGCCCACTGCCACCATCAGCA	1140
QY	1141	CGCTAAACGGAAGGAGACTTCGGATATAACAGCAGGAGCTGGCAACAGAAAAACA	1200
DB	1141	CGCTAAACGGAAGGAGACTTCGGATATAACAGCAGGAGCTGGCAACAGAAAAACA	1200
QY	1201	AAGGCTGTGTGGAAGAAATAATCTTTGGATGCTGATGTACAACTACGACTACGAAAG	1260
DB	1201	AAGGCTGTGTGGAAGAAATAATCTTTGGATGCTGATGTACAACTACGACTACGAAAG	1260
QY	1261	CAGCGTGGCTGATCAGTTCGCTTAGATTTAATTCCTGTTTAACTCAGAGCTTT	1320
DB	1261	CAGCGTGGCTGATCAGTTCGCTTAGATTTAATTCCTGTTTAACTCAGAGCTTT	1320
QY	1321	GGTAAAGTTTCTCATGTTTCAAGCTGGGTAAAGTTAGTTGTTTGAAGAGATTGGTG	1380
DB	1321	GGTAAAGTTTCTCATGTTTCAAGCTGGGTAAAGTTAGTTGTTTGAAGAGATTGGTG	1380
QY	1381	TACCAAGTAAACAACTTATCCTCTTTTAACTCTTTTGTCTTTGAAGTAAAAAAA	1440
DB	1381	TACCAAGTAAACAACTTATCCTCTTTTAACTCTTTTGTCTTTGAAGTAAAAAAA	1440
QY	1441	AAAAAATAAAAAAATAAAAAA 1463	
DB	1441	AAAAAATAAAAAAATAAAAAA 1463	

RESULT 2

AAA27926

ID AAA27926 standard; cDNA; 1478 BP.

XX AAA27926;

AC AAA27926;

DT 12-SEP-2000 (first entry)

XX Maize Rad2/FEN-1 cDNA.

XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.

XX	Zea mays.
OS	
XX	
XX	Location/Qualifiers
KY	97..1236
FH	/*tag= a
FT	
ET	
XX	
PN	WO200036109-AI.
PD	22-JUN-2000.
XX	
XX	16-NOV-1999; 99WO-US27147.
PR	15-DEC-1998; 98US-0112332.
XX	(PTON-) PIONEER HI-BRED INT INC.
PA	Mahajan PB;
XX	
PI	WPI; 2000-452026/39.
DR	P-PSDB; AAY95310.
XX	
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants -
XX	Example 1; Page 77-79; 85pp; English.
CC	The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AAY95310). The cDNA was isolated from a library prepared from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the Rad2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving plants. RAD2/FEN-1 endonucleolytic activity is essential in DNA replication and nucleotide excision and repair reactions. . The exolytic activity is involved in double strand break repair and end joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.
XX	
SQ	Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;
	Query Match 97.2%; Score 1421.8; DB 21; Length 1478;
	Best Local Similarity 98.5%; Pred. No. 0;
	Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps
QY	7 ATATAGTCGCGGCCGGTCTTTCTGCGGCACATCCGGCTCAGCCGCGCGCCGACCACCGGCC 66
Db	19 ATATAGTCGCGGTCCGGTCTTTCTGCGGCACATCCGGCTCAGCCGCGCGCCGACCACCGGCC 78
QY	67 ACAGCGCCGCGACGAGCATGGGCATCAAGGGTTTGAGAAACTGCTGGCGGACAATGCG 126
Db	79 ACAGCGCCGCGACGAGCATGGGCATCAAGGGTTTGAGAAACTGCTGGCGGACAATGCG 138
QY	127 CCCAAGGCGATGAAGGACGAGAAGTTTCAGAGACTACTTCGGCCGCCAAAATGCGCGTCGAC 186
Db	139 CCCAAGGCGATGAAGGACGAGAAGTTTCAGAGACTACTTCGGCCGCCAAAATGCGCGTCGAC 198
QY	187 GCAGCATGAGCATATACAGTTCCTCTGATTGTAGTTGAAGGACGAGCATGGAAACTCTC 246
Db	199 GCCAGCATGAGCATCTACAGTTCCTCTGATTGTAGTTGAAGGACGAGCATGGAAACTCTC 258
QY	247 ACAATGAAGCTGGTGAAGTCACTAGTTCATTTCGAAGGAATGTTTCAACCGGACAATAAGA 306
Db	259 ACAATGAAGCTGGTGAAGTCACTAGTTCATTTCGAAGGAATGTTTCAACCGGACAATAAGA 318

Db 1399 GTAAACAAACTTATGCTGTTTTTACTTCTTGTCTTGAAGTATGATGCCAGTAAA 1458
 QY 1447 AAAAAAAAAAAAAA 1463
 Db 1459 AAAAAAAAAAAAAA 1475

RESULT 3
 ID AAA27924 standard; cDNA; 1541 BP.
 AC AAA27924;
 DT 12-SEP-2000 (first entry)
 DE Maize Rad2/FEN-1 cDNA.
 KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
 endonuclease; exonuclease; DNA repair; gene targeting; ss.
 OS Zea mays.
 XX Key Location/Qualifiers
 FT CDS 79..1218
 FT /*tag= a
 PN WO200036109-A1.
 XX 22-JUN-2000.
 XX 16-NOV-1999; 99WO-US27147.
 XX 15-DEC-1998; 98US-0112332.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Mahajan PB;
 XX WPI; 2000-452026/39.
 XX P-PSDB; AAY95308.
 XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
 recombination and repair in transgenic plants, e.g. for gene targeting
 and the production of male sterile plants -
 Example 1; Page 73-74; 85pp; English.
 XX The present sequence is that of maize cDNA (ATCC PTA-533) coding
 for RAD2/FEN-1 (see AAY95308). The corresponding RNA was isolated
 from a B73 line seedling after a 10 day drought, heat shocked for
 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1
 is a structure specific endonuclease which under certain conditions
 also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used
 to produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
 The protein is involved in the regulation of DNA repair and
 recombination in plant systems and therefore may be used for
 improving gene targeting during further recombinant DNA protocols
 involving plants. RAD2/FEN-1 endonucleolytic activity is essential
 in DNA replication and nucleotide excision and repair reactions.
 The exolytic activity is involved in double strand break repair and
 end joining. The protein is also useful in strand exchange
 reactions during homologous recombination. These functions may be
 useful in gene targeting and in the production of male sterile
 plants. The efficacy of gene targeting can be improved by the
 overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 be produced by the down regulation of Rad2/FEN-1 expression.
 XX Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other;
 Query Match 96.4%; Score 1409.8; DB 21; Length 1541;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CGCGGTTTCTTGGCCACTCCGCTCAGCGCGCGCCGACCCGACAGCGCGCA 78
 Db 13 CGGCGTTTCTTGGCCACTCCGCTCAGCGCGCGCCGACCCGCGCGCA 72
 QY 79 GACGAGATGGCATCAAGGGTTTGACGAACTGCTGGCGGACAATGGCCCAAGGCGATG 138
 Db 73 GACGAGATGGCATCAAGGGTTTGACGAACTGCTGGCGGACAATGGCCCAAGGCGATG 132
 QY 139 AAGGAGCAGAAAGTTTCGAGAGCTACTTCGGCGGCAAAATCCCGTCGACGCCACATGAGC 198
 Db 133 AAGGAGCAGAAAGTTTCGAGAGCTACTTCGGCGGCAAAATCCCGTCGACGCCACATGAGC 192
 QY 199 ATATACCAAGTTCTCTGATTTGTTGGAAGGACAGGCACTGGAACCTCTCAAAATGAAGCT 258
 Db 193 ATATACCAAGTTCTCTGATTTGTTGGAAGGACAGGCACTGGAACCTCTCAAAATGAAGCT 252
 QY 259 GGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTGGAGCG 318
 Db 253 GGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTGGAGCG 312
 QY 319 GGAATCAAGCCAGTTTATGTTTGTGCAAGCCCTCTGATATGAAGAAACAAGAGCTT 378
 Db 313 GGAATCAAGCCAGTTTATGTTTGTGCAAGCCCTCTGATATGAAGAAACAAGAGCTT 372
 QY 379 GCTAAAGATATCAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGAGTAJAGTA 438
 Db 373 GCTAAAGATATCAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGAGTAGAGTA 432
 QY 439 GGAGATAAGATGCGATTGAAAAATGAGCAAGAGGACTGTAAAGGTCAACAAGCAACAC 498
 Db 433 GGAGATAAGATGCGATTGAAAAATGAGCAAGAGGACTGTAAAGGTCAACAAGCAACAC 492
 QY 499 AACGAAGATTGTAACGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGCACTTCT 558
 Db 493 AACGAAGATTGTAACGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGCACTTCT 552
 QY 559 GAAGCAGAGCAGAAATGTCAGCCCTTTGCTATAACGATAAGGTGTTGCTGTGTTCA 618
 Db 553 GAAGCAGAGCAGAAATGTCAGCCCTTTGCTATAACGATAAGGTGTTGCTGTGTTCA 612
 QY 619 GAAGATATGGATCCCTTACTTTTGGGGCTCCAGGTTCCCTGCTCATTTAATGGATCA 678
 Db 613 GAAGATATGGATCCCTTACTTTTGGGGCTCCAGGTTCCCTGCTCATTTAATGGATCA 672
 QY 679 AGTTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTGGAGGAGCTTGA 738
 Db 673 AGTTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTGGAGGAGCTTGA 732
 QY 739 CTCACCATGGACAGTTTCATGATTGTCGATCTGTGATGTCGACTATTGTGATAGC 798
 Db 733 CTCACCATGGACAGTTTCATGATTGTCGATCTGTGATGTCGACTATTGTGATAGC 792
 QY 799 ATCAAAGTATCGGGGGGCAACACAGCTCTGAAACTTATTGTCGAACATGGGTCATAGAA 858
 Db 793 ATCAAAGTATCGGGGGGCAACACAGCTCTGAAACTTATTGTCGAACATGGGTCATAGAA 852
 QY 859 AGCATCTGGAGATCTTAATAAGACAGATATAAATCTCTGACGACTGGCTTACCAA 918
 Db 853 AGCATCTGGAGATCTTAATAAGACAGATATAAATCTCTGACGACTGGCTTACCAA 912
 QY 919 GAAGCTCGACCTTGTGTTCAAGGAGCCCTTAATGTGCACATTTGATATTCCTGAGCTAAATGG 978
 Db 913 GAAGCTCGACCTTGTGTTCAAGGAGCCCTTAATGTGCACATTTGATATTCCTGAGCTAAATGG 972
 QY 979 ACTGCACCTGTAGAGGAGGCTCATAGTTTCTCGTAAAGATAATGTTTCAACGAA 1038
 Db 973 ACTGCACCTGTAGAGGAGGCTCATAGTTTCTCGTAAAGATAATGTTTCAACGAA 1032
 QY 1039 GATCGGTTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAATAAATCGTCGCAAGA 1098
 Db 1033 GATCGGTTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAATAAATCGTCGCAAGA 1092

QY 1099 AGACTCGAGTCTCTTTTCAAGCCAACTGCCACCATCAGCACCCCTAAAACGGAAGGAG 1158
 DB 1093 AGACTCGAGTCTCTTTTCAAGCCAACTGCCACCATCAGCACCCCTAAAACGGAAGGAG 1152
 QY 1159 ACTTCGGGATAAACAAGCAGCTCGGACAAAGAAACAAAGGCTGGTGGAAAGAG 1218
 DB 1153 ACTTCGGGATAAACAAGCAGCTCGGACAAAGAAACAAAGGCTGGTGGAAAGAG 1212
 QY 1219 AAATAATCTTGGATCTTGTATGATACAACTAGACTACGAAAGCAGCGGTGGCGTGATCAC 1278
 DB 1213 AAATAATCTTGGATCTTGTATGATACAACTAGACTACGAAAGCAGCGGTGGCGTGATCAC 1272
 QY 1279 TFCGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGTCTATG 1338
 DB 1273 TFCGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGTCTATG 1332
 QY 1339 TTTCAGCTGGGTAAAGTAGTGTGTTTGAAGAGATTGGTGTACCAAGTACAAACTT 1398
 DB 1333 TTTCAGCTGGGTAAAGTAGTGTGTTTGAAGAGATTGGTGTACCAAGTACAAACTT 1392
 QY 1399 ATCGCTGTTTTTACTTCTTGTCTCTTTTGAAGTA 1431
 DB 1393 ATCGCTGTTTTTACTTCTTGTCTCTTTGAAGTA 1425

RESULT 4
 ID AAA27925 standard; cDNA; 1381 BP.
 XX AAA27925;
 AC AAA27925;
 DT 12-SEP-2000 (first entry)
 DE Maize Rad2/FEN-1 cDNA.
 KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
 endonuclease; exonuclease; DNA repair; gene targeting; ss.
 OS Zea mays.
 FH Key Location/Qualifiers
 FT CDS 37..1176
 FT /*tag= a
 XX
 PN WO200036109-A1.
 XX
 PD 22-JUN-2000.
 XX
 PF 16-NOV-1999; 99WO-US27147.
 XX
 PR 15-DEC-1998; 98US-0112332.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Mahajan PB;
 XX
 WPI: 2000-452026/39.
 DR P-PSDB; AAY95309.
 XX
 PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
 recombination and repair in transgenic plants, e.g. for gene targeting
 and the production of male sterile plants -
 XX
 PS Example 1; Page 75-76; 85pp; English.
 XX
 CC The present sequence is that of maize cDNA coding for RAD2/FEN-1
 (see AAY95309). The cDNA was isolated from a cDNA library prepared
 from W23-line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a
 CC structure specific endonuclease which under certain conditions also
 acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
 CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
 CC The protein is involved in the regulation of DNA repair and
 recombination in plant systems and therefore may be used for

CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 XX
 SQ Sequence 1381 BP; 441 A; 269 C; 346 G; 325 T; 0 other;
 Query Match 92.3%; Score 1350.2; DB 21; Length 1381;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 49 CGCCGCCGCCCGCCACAGCCGCGCAGAGATGGCATCAAGGTTTGACGAAA 108
 DB 1 CGACCCACGCGTCCGCGCCACAGCCGCGCAGAGATGGCATCAAGGTTTGACGAAA 60
 QY 109 CTGCTGGCGGACAAATCGCCCAAGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGCG 168
 DB 61 CTGCTGGCGGACAAATCGCCCAAGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGCG 120
 QY 169 CGCAAAATCGCGTCGACGCCAGCATGACCATATACCAGTTCTCTGATTAGTTGGAAGG 228
 DB 121 CGCAAAATCGCGTCGACGCCAGCATGACCATATACCAGTTCTCTGATTAGTTGGAAGG 180
 QY 229 ACAGGATGGAACCTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTATTTCGAATG 288
 DB 181 ACAGGATGGAACCTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTATTTCGAATG 240
 QY 289 TTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGGC 348
 DB 241 TTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGGC 300
 QY 349 AGCCCTCTGATATGAAGAAACAAGAGCTTGTAAAGATACCTCAAAACAGATGATGCA 408
 DB 301 AAGCCCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATACCTCAAAACAGATGATGCA 360
 QY 409 ACCAAGATCTGACTGAGGCGAGTAGAGTAGGAGATAAGATCGGATTGAAAATTTGAGC 468
 DB 361 ACCAAGATCTGACTGAGGCGAGTAGAGTAGGAGATAAGATCGGATTGAAAATTTGAGC 420
 QY 469 AAGAGGACTGTAAAGTCAACAGGCAACACACAGAGATTGTAACGGCTATTAAAGACTT 528
 DB 421 AAGAGGACTGTAAAGTCAACAGGCAACACACAGAGATTGTAACGGCTATTAAAGACTT 480
 QY 529 ATGGGGTTCCTGTTTAGAGGCACCTTCTGAAGCAGACAGCAANTGTGAGCCCTTTGC 588
 DB 481 ATGGGGTTCCTGTTTAGAGGCACCTTCTGAAGCAGACAGCAANTGTGAGCCCTTTGC 540
 QY 589 ATAAAGATAAGGTGTCGCTGTTGCTTCAAGAGATATGGACTCCCTTACTTTTGGGCT 648
 DB 541 ATAAAGATAAGGTGTCGCTGTTGCTTCAAGAGATATGGACTCCCTTACTTTTGGGCT 600
 QY 649 CCACGGTTCCTTCCTCATTTAATGGATCCAAAGTTTCCAAGAAAATACCTGTGATGGAATTT 708
 DB 601 CCACGGTTCCTTCCTCATTTAATGGATCCAAAGTTTCCAAGAAAATACCTGTGATGGAATTT 660
 QY 709 GATGTTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGACCAAGTTCATTGATTGTCG 768
 DB 661 GATGTTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGACCAAGTTCATTGATTGTCG 720
 QY 769 ATCCTGTGTGGATGTGACTATTCTGTAGCATCAAGGATATCGGGGCGCAACAGCTCTG 828
 DB 721 ATCCTGTGTGGATGTGACTATTGTGTAGCATCAAGGATATCGGGGCGCAACAGCTCTG 780
 QY 829 AAACCTATTCTCAACATGGGTCCATAGAAGCATCTTTGGAGATCTTTAATAAACAACA 888
 DB 781 AAACCTATTCTCAACATGGGTCCATAGAAGCATCTTTGGAGATCTTTAATAAACAACA 840

QY 925 CGACGCTGTTCAGGAGCCTAATG---TCACATTTGGATATTCCTGAGCTAAATGGACT 981
 Db 838 CACCAGCTCTTCTTGGAACTGAGTCTGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 897
 QY 982 GCACCTGATGAGGAGGCTCATAAGTTTCTGTGTAAGAGATAATGTTTCAACGAAGAT 1041
 Db 898 GAGCCAAATGAAGAAGAGCTGATCAAGTTTCATGTGTGTGTAAGAGAGCTTCTCTCAGGAG 957
 QY 1042 CGGTGTACAAGGCCATAGAGAAGATCAAAATCTGCCAAGATTAATCGTCCCAAGGAAGA 1101
 Db 958 CGAATCCGAGTGGGTCAAGAGCTGAGTAAGAGCCGCCAAGGAGCACCAGGCGCGC 1017
 QY 1102 CTCGAGTCTCTTTTCAAGCCAACTGCCACACATCAGCACC 1142
 Db 1018 CTGGATGATTCTTCAAGGTGACCGGCTCACTCTCTTTCAGC 1058

RESULT 6

AA02107
 ID AAX02107 standard; cDNA; 1144 BP.

AC AAX02107;

XX 23-APR-1999 (first entry)

DE Human FEN-1 cDNA.

XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasia; antineoplastic agent; cleavage; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT 1..1144
 CDS /tag= a
 FT /product= "FEN-1"

XX US5874283-A.

XX 23-FEB-1999.

XX 30-MAY-1995; 95US-0455968.

XX 30-MAY-1995; 95US-0455968.

XX (HARR/) HARRINGTON J J.

XX (HSIE/) HSIEH C.

XX (LIEB/) LIEBER M R.

XX Harrington JJ, Hsieh C, Lieber MR;

XX WPI: 1999-179985/15.

XX P-PSDB; AAW92504.

XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides

XX Disclosure; Fig 1B; 58pp; English.

XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.

XX Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;

XX Query Match 20.7%; Score 302.6; DB 20; Length 1144;

Best Local Similarity 56.6%; Pred. No. 1.1e-66;
 Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;
 QY 85 ATGGGATCAAGGGTTTGGAGAACTGTGGCGGACAAATCGGCCAAGGCGATGAAGAG 144
 Db 1 ATGGGAAATTCAGAGGCTGGCCAACTAAATGCTGTATGTGGCCCGCAGTGCATCGGAG 60
 QY 145 CAGAAGTTTCAGAGCTTCTGGCCGCAAAATCGCGTCGACGCCAGCATGAGCATATAC 204
 Db 61 AATGACATCAAGAGCTTCTTGGCCGTAGGTGCGCATGATGCCTCTATGAGCTTTAT 120
 QY 205 CAGTTCTCTGATTGTAGTTGGAGGACAGCATGGAACCTCTCACAAATGAAGCTTGTAA 264
 Db 121 CAGTTCTCTGATTGCTTTCGCCAG---GCTGGGATGTGTCGAGAATGAGGAGGTGAG 177
 QY 265 GTCACCTAGTCATTTGCAAGGAATTTCAACCCGCAATATAGATTACTTGGAAAGCGGATC 324
 Db 178 ACCACAGCCACCTGATGGCATTTCTACCCGACCATTCGCATGATGGTAGACGGGATC 237
 QY 325 AAGCCAGTTTATGTTTGTATGGCAAGCCTCTCTGATATGAAGAAACAAGAGCTTGTCTAA 384
 Db 238 AAGCCGCTGATGCTTGTGTCAGAGCCGCCACAGCTCAAGTCAGCGGAGCTGCCCAA 297
 QY 385 AGATATCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGATGAGGTAGGAGAT 444
 Db 298 CGCAGTGAGCGCGGCTGAGGCAGAGAGCAGCTGCAGCAGGCTCAGGCTGTGGGCGC 357
 QY 445 AAGATGCGATTCAAAATTTGACCAAGAGACTGTTAAAGTTCACAGGCAACACAGCA 504
 Db 358 GAGCAGGAGGTGGAATAATTAAGCGGCTGTTGAAGGTCACTAAGCAGCACAATGAT 417
 QY 505 GATTGTAACCGGCTATTAGACTTATGGGGTTCCTTGTAGAGGACCTTCTTGAAGCA 564
 Db 418 GAGTCAACATCTGCTGAGCCTCATGGGCATCCCTATCTTGTATGACCCAGTGGCA 477
 QY 565 GAAGCAGATGTCAGCCCTTTGTCATAAAGTAAAGTGTTCGCTTGTGCTTCAAGAGAT 624
 Db 478 GAGGCCAGCTGTGCTGCCCTGTTGTAAGGCTGGCAAGTCTATGCTGCGGTATCAGGAG 537
 QY 625 ATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCTGCTATTTAATGGATCCAAGTCC 684
 Db 538 ATGGACTCCCTTACTTTCGGCAGCCCTGTCTAATGGCAGACCTGATGCTGAGTGAAGCC 597
 QY 685 AAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTGGAGGAGCTTGAACCTACC 744
 Db 598 AAAAGCTGCCAATCCAGGAATTCACCTGAGCGGATTCGACGAGCTGGGCTGAC 657
 QY 745 ATGGACCACTTCATTTGATTTGTCATCCTCTGTGGATGTGACTATTGTATAGATCAAA 804
 Db 658 CAGGAACAGTTTGTGGATCTGTGCATCTCTAGGCAAGTACTACTGTGAGAGTATCCGG 717
 QY 805 GGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGTCCATAGAAGCATC 864
 Db 718 GTATTTGGCCCAAGCGGCTGTGGACCTCATCCAGAAGCAGCAGAGCATCGAGGAGAT 777
 QY 865 TTGGAGAACTTAAATAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCT 924
 Db 778 GTGGCGGACTTGACCCCAACAGTACCTGTGCCAGAAAATTTGGCTCCACAGAGGCT 837
 QY 925 CGACGCTTCTTCAAGGAGCCTAATG---TCACATTTGGATATTCCTGAGCTTAAATGGACT 981
 Db 838 CACCAGCTCTTCTTGAACCTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAATGGAGC 897
 QY 982 GCACCTGATGAGGAGGCTCATAAGTTTCTGTTAAAGATAATGTTTCAAGAGAT 1041
 Db 898 GAGCCAAATGAAGAAGAGCTGATCAAGTTTCATGTGTGGTGAAGAGAGCTTCTCTGAGGAG 957
 QY 1042 CGGTGTACAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCCCAAGGAAGA 1101
 Db 958 CGAATCCGAGTGGGTCAAGAGCTGAGTAAGAGCCGCCAAGCAGCACCAGGCGCGC 1017
 QY 1102 CTCGAGTCTCTTTTCAAGGCCAACTGCCACACATCAGCACC 1142

Db 1018 CTGGATGATTCTTCAAGGTGACCGGCTCACTCTCTTCAGC 1058

RESULT 7
AA02111

ID AAX02111 standard; DNA; 2033 BP.

XX AC AAX02111;

XX 23-APR-1999 (first entry)

XX Human FEN-1 genomic DNA.

XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasia; antineoplastic agent; cleavage; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 104..1240

FT /*tag= a
FT /product= "FEN-1"

XX US5874283-A.

XX 23-FEB-1999.

XX 30-MAY-1995; 9505-0455968.

XX 30-MAY-1995; 9505-0455968.

XX (HARR/) HARRINGTON J J.

XX (HSIE/) HSIEH C.

XX (LIEB/) LIEBER M R.

XX Harrington JJ, Hsieh C, Lieber MR;
XX WPI; 1999-179985/15.

XX P-PSDB; AAW92508.

XX DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides

XX Disclosure; Fig 5A-B; 58pp; English.

XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.

XX Sequence 2033 BP; 500 A; 493 C; 579 G; 461 T; 0 other;

Query Match 18.9%; Score 276.6; DB 20; Length 2033;
Best Local Similarity 55.9%; Pred. No. 7.9e-60;
Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGATCAAGGTTTTCAGCAAACTCTGGCGGACAAATCGCCCAAGCGATGAAGAG 144

Db 104 ATGGAATTCACGCGCTTGCCAACTAATTCCTGATGTGGCCCGCCAGTCCGTGAG 163

QY 145 CAGAAATTCAGAGCTACTTCGGCGCAAAATCGCCGTCAGCCAGCATGACATATAC 204

Db 164 AATGACATCAAGAGCTACTTTGGTCGCAAAAGTGCCATCGATGCCCTCCATGAGCATCTAC 223

QY 205 CAGTCTCTGATTGTGGAAGACAGCATGGAACACTCTCACAATGAAGCTGGTGAA 264

Db 224 CAGTCTCTGATTGTGTCGTAG--GGTGGGATGTGTCGACAGACGAGGAGGTGAG 280

QY 265 GTCCTAGTCAATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTTGGAGCGGGAATC 324
Db 281 ACCACGAGCC---TGATGGGCATGTTCTACCGTACCATGCG---CATGGAGAATGGCATC 334

QY 325 AAGCCAGTTTATGTTTGTGATGGCAAGCCCTCCTCATATGAAGAAACAAGAGCTTGTCTAAA 384

Db 335 AAGCCTGTGTACGCTTTTGTATGGCAAAACCCACAGCTGAAGTCAAGCGAGCTGCCCAAG 394

QY 385 AGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGCAGTAGAGTAGGAGAT 444

Db 395 CGCAGTCAGAGCGCGCGGAGGCTGAGAAGCACTGACAGAGCTCAGCAGGCTGGGATG 454

QY 445 AAAGATCGGATTGAAAATTTGAGCAAGAGGACTGTAAGGTCACAAAGGCAACAACGAA 504

Db 455 GAGGAGGAGGTGGAGAAGTTTCAACCAAGAGGCTCGTAAGGTCACCAAGCAACAATGAT 514

QY 505 GATTGTAAACGGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCTGAAGCA 564

Db 515 GAGTGCAAACACCTGCTGAGCCTCATGGGATCCCTTACCTTGATGACCCAGCGAGGCA 574

QY 565 GAAGCAGAATGTGCAGCCCTTTGCATAAAGATGTTGCGCTGTTGCTTCAAGAAGAT 624

Db 575 GAGCCAGCTGTGCTGCCCTGGCAAGGCTGGCAAGCTCTATGCTGGCGCCACGAGGAC 634

QY 625 ATGACTCCCTTACTTTTGGGGTCCACGGTTCCTTCGTCATTTAATGGATGCCAAGTTC 684

Db 635 ATGACTGCCCTCACTTTTGGCAGCCCGCTGCTAATGCGACACTTAAGTGCAGTGAGGCC 694

QY 685 AAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGAACCTACC 744

Db 695 AAGAGCTGCCCATCCCAAGAGTTCCTATCTGAGCGCGCTCCTGACAGGAGCTGGGTCTGAAC 754

QY 745 ATGACCAGTTCAATTTGTCATCCTGTGTGGATGTGACTATTGTATAGCATCAAA 804

Db 755 CAGGAGCAGTTTCTGGATCTGTCATCCTGCTGGTAGCGACTACTCGGAGAGATCCGT 814

QY 805 GGTATCGGGGGCAACAGCTCTGAAACTTATTGTCGTCACATGGGTCCATAGAAGCATC 864

Db 815 GGCATTGGCCCAAGCGGCTGTGGATCTCATCCAGAAACATAAGAGCATCGAGAGATC 874

QY 865 TTGGAGAATCTTAATAAGACAGATATCAAAATTCCTCAGGAGCTGGCCTTACCAGAAGACT 924

Db 875 GTGAGCGGCTGGACCCCAAGTACCCCGTTCCAGAGAACTGGCTCCACAGGAAGGCC 934

QY 925 CGAGCTCTGTTCAAGGAGCCCTAA---TGTCACATTTGGATATTCCTGAGCTAAAATGGACT 981

Db 935 CAGCAGCTCTTCTCTGGAGCCAGAAGTAGTGACCCAGAGCTGTGGAGCTGAAGTGGAGC 994

QY 982 GCACCTGATGAGGAGGCTCATTAAGTTTCTGTTGTAAGATATATGTTTCAACGAGAT 1041

Db 995 GAGCCAAATGAAGAAGAGTTGGTCAAAATTTATGTGTGTGAAAAGACAGTTTGTGAAAG 1054

QY 1042 CGGTGTCAAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAATAAATCGTCGCAAGGAAGA 1101

Db 1055 CGAATTCGAGTGGGTCAAGCGGCTGAGTAAGAGCCGCCAGGCGACACCCAGGAGGC 1114

QY 1102 CTCGAGTCTCTTTTCAAGCCCAACTGCCACCATCATGACACCGCTTAAACGAGGAGACT 1161

Db 1115 CTCGATGATTCTTCAAGGTGACAGGCTCACTCTCCTCAGC-----TAAGCGC 1162

QY 1162 TCGGATAAACAAGCAGAGGAGCTGCGAACAAGAAACAAGAGCTGTGTGAAGAAGAAA 1221

Db 1163 AAGSAGCAGAACCCCAAGGGGCTGCTAAGAAGAAACAAGAGCTGTGGGGAGCGGGAAG 1222

QY 1222 T 1222

Db 1223 T 1223

RESULT 8

ABL14287

ID ABL14287 standard; cDNA; 1300 BP.

PR 30-MAY-1995; 95US-0455968.
 XX (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX WPI; 1999-179985/15.
 DR P-PSDB; AAW92505.
 XX
 XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 XX Disclosure; Fig 2B; 58pp; English.
 XX
 CC This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 XX
 SQ Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other;
 Query Match 18.6%; Score 272.2; DB 20; Length 1930;
 Best Local Similarity 55.1%; Pred. No. 1e-58;
 Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;
 QY 85 ATGGGATCAAGGTTTGAAGAACTCTGGCGGACAAATCGCCGAGCGATGAAGAG 144
 DB 1 ATGGGAATTCACGGCTTGCCAACTAATGTCTGATGTGGCCGCCAGTGCCTCGTG 60
 QY 145 CAGAAGTCGAGACTCTCGCGCAAAATCGCGTCGACCGCAGCATGACATATAC 204
 DB 61 AATGACATCAAGAGCTACTTGTGTGAAGTGGCCATCGATCGCTCCATGAGCATCTAC 120
 QY 205 CAGTCTCTGATTTAGTGAAGGACAGGCATGGAACTCTCAAAATGAAGCTGGTAA 264
 DB 121 CAGTCTCTGATTTCTGTCTGAG---GGTGGGATGTGCTGCAGAACAGGAGGTGAG 177
 QY 265 GTCACTAGTCAATTTGAAGAAATGTTCAACCGGACAAATAGATTACTGGAAGCGGAATC 324
 DB 178 ACCACGAGCTGTATGGGCATGTTAT-----GGCAACCATCCGATGGAGAATGGCATC 231
 QY 325 AAGCCAGTTTATGTTTGTATGCAAGCTCTCTGATATGAAGAAACAGAGCTTGCTAA 384
 DB 232 AAGCCGTGTACGCTTTGATGCAACACCAACACAGCTGAAGTCAAGCGAGCTGGCCAG 291
 QY 385 AGATACTCAAAAAGAGATGATGCAACCAAGATPCTGACTGAGGCAGTAGAGTAGGAGAT 444
 DB 292 CGCAGTGAGAGGCGCCGAGGCTGAGAAGCACTCAGCAGGCTCAGGAGGTGGGATG 351
 QY 445 AAGATGCGATGCAAAATTTGACAGAGAGACTGTAAAGTCAACAGGCAACACACGAA 504
 DB 352 GAGGAGGAGGTGGAGAAGTTCACCAAGAGGCTGTGAAGGTCAACCAAGCAACAAATGAT 411
 QY 505 GATTGTAAACGGCTATTAAAGCTATTATGGGGTTTCCTGTTGTAGAGCAGCTTCTGAAGA 564
 DB 412 GAGTGAACACCTCTGTGAGCTCTATGGGATCCCTTACCTTGATGCACCCAGCGAGGA 471
 QY 565 GAAGCAAGATGTGACGCCCTTTGCATAAAGATGAAGGTGTGCTGTGCTTCAAGAAGAT 624
 DB 472 GAGGCCAGCTGTGCTGCCCTTGGCAAGGCTGCAAAAGTCTATGCTCGCCGACGAGGAC 531
 QY 625 ATGGACTCCCTTACTTTTGGGGCTCCACGGTCTCTTCGTCATTTAATGATCAAGTCC 684
 DB 532 ATGGACTGCTCTACTTTTGGCAGCCCGCTGCTAATGCGACACTTAACTGCCAGTGAGGCC 591

QY 685 AAGAAAATACCTGTGTATGGAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCACC 744
 DB 592 AAGAGCTGCCATCCCAAGAGTTCCATCTGAGCGGCTCTCGCAGGAGCTGGGTCTCAAC 651
 QY 745 ATGACCAAGTTCAATTTGATTTGTCATCTGTGTGGATGTGACTATTGTATAGATCAAA 804
 DB 652 CAGGAGCAGTTTGTGGATCTGTGCATCTGCTGGTAGCAGCTACTCGGAGAGCATCCGT 711
 QY 805 GGTATCGGGGGCAACAGCTCTGAACATTTATTCGTCAACATGGTCCATAGAAGCATC 864
 DB 712 GGATTTGGCCCAAGCGGGCTGTGGATCTCATCCAGAAACATAGAGCATCTGAGGAGATC 771
 QY 865 TTGGAGAATCTTAATAAGACAGATATCAAAATCTCTGAGACTGGCCTTACCAAGAGCT 924
 DB 772 GTGAGCGGCTGGACCCCAAGTACCCCGTCCAGAGAACTGGCTCCACAAGGAAGCC 831
 QY 925 CGAGCTGTTCAGAGGAGCCTAA---TGTCACATTTGGATATTCCTGAGCTAAATGGACT 981
 DB 832 CAGCAGCTCTTCTGGAGCCAGAGTAGTAGGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 891
 QY 982 GCACCTGATCAGGAGGCTCTCAATAGTTTCTCTGTAAAGATATGTTTCAACGAAGAT 1041
 DB 892 GAGCCAAATGAAGAGAGTTGGTCAAAATTTATGTGGTGAAGAGAGTTTCTGAAAGAG 951
 QY 1042 CGGTGACAAAGGCCATAGAGAAAGATCAAAATCTGCCAAGAAATAAATCGTCGCAAGGA 1101
 DB 952 CGAATTCGCAAGTGGGTCAAGCGGCTGAGTAAGAGCGCCAGGCGAGCACCCAGGGAGC 1011
 QY 1102 CTGAGTCTCTTTTCAAGCAACTGCCACCACATCAGCACCGCTAAACGGAGGAGACT 1161
 DB 1012 CTCGATGATTTCTTCAAGGTGACAGGCTCACTCTCTCAGC-----TAAGCGC 1059
 QY 1162 TCGGATAAACAACAGGAGGCTGGCAACAAGGCTGGTGAAGGCTGGTGAAGAGAGAA 1221
 DB 1060 AAGGAGCAGAACCCAGGGGCTGCTAAGAGCAAGCAAGAGACTGGGGGAGCGGGAAG 1119
 QY 1222 T 1222
 DB 1120 T 1120
 RESULT 10
 AAX02109
 ID AAX02109 standard; cdna; 1149 BP.
 XX
 AC AAX02109;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Yeast FEN-1 cdna.
 XX
 KW FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasia; antineoplastic agent; cleavage; ss.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PF 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX WPI; 1999-179985/15.
 DR P-PSDB; AAW92506.

PT DNA encoding flap endonuclease polypeptides - useful for producing
 XX e.g. recombinant polypeptides
 PS Disclosure; Fig 3B; 58pp; English.

XX This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.

XX Sequence 1149 BP; 410 A; 182 C; 279 G; 278 T; 0 other;

Query Match 18.5%; Score 270.2; DB 20; Length 1149;
 Best Local Similarity 55.5%; Pred. No. 2.6e-58;
 Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

QY 85 ATGGGATCAAGGTTTGACGAACACTCTGGCGACAAATCGCCCAAGCGCATGAAGAG 144
 DB 1 ATGGGATTAAGGTTTGAATGCAATATATATCGGAACATGTTCCCTCTCTATCAGGAA 60
 QY 145 CAGAAGTTCGAGAGCTACTTTCGGCGCAAAATCGCGTCGAGCGCAGCATGATATAC 204
 DB 61 AGGATATCAAGAGCTTTTGGCAGAAAGTTGCCATCGATGCTCTATCTCTATAT 120
 QY 205 CAGTTCCTGATTTAGTGAAGGACAGCGATGGAACCTCTCACAATGAAGCTGGTGA 264
 DB 121 CAGTTTTTAATGCTGAAGACAGCAAGCGTGGCGAGTTGACCAATGAAGCGGTGA 180
 QY 265 GTCACTAGTCATTTCGAAGAAATGTTCAACGGGACATATAGATTACTGGAAGCGGATC 324
 DB 181 ACAAGCTCACACTTGATGGGTATTTTATAGGACACTGAGAATGATTGATACGGTATC 240
 QY 325 AAGCAGCTTTATGTTTGTGGAAGCGCTCTCTGATATGAAGAACAGAGCTTGCTAAA 384
 DB 241 AAGCCTTTGATGTTTCGAGCGCAACCTCCAGCTTTGAAATCTCATGATGACAAAG 300
 QY 385 AGATACCTCAAAAGAGATGATGACCAACAGATCTGACTGAGGCGAGTAGAGTAGAGAT 444
 DB 301 CGGTCTTCAAGAGGCTGGAACAGCAAGAAAGAACTGGCAGAGGCAACACAGAAATGGA 360
 QY 445 AAGATCGGATTTGAATAATTTGAGCAAGAGGACTGTAAGGTCACAAAGCCACACACAA 504
 DB 361 AAGATGAAGCAAGAAAGA-----AGATTGTTGAAGGTTCTCAAAAGAGCATTAATGA 411
 QY 505 GATTGTAACGGCTATTAAAGACTTATGGGGTTCTCTGTTGAGAGGACCTTCTGAAGCA 564
 DB 412 GAAGCCCAAAATTTACTAGACATTAATGGAAATCCATATATAATAGCCCAACGGAAGCT 471
 QY 565 GAAGCAGAATGTGAGCCCTTTTGCAATAAAGATGAGTGTGCTGTTGCTTTCAGAAAGAT 624
 DB 472 GAGGCTCAATGTGCTGAGTTGGCAAGAAAGGAAAGGTGATGCGCAGCAAGTGAAGAT 531
 QY 625 ATGAGCTCCCTTACTTTTGGGCTCCAGGTTCTCTGCTCATTTAATGGATCCAGTTCC 684
 DB 532 ATGGACACACTGTTTATAGAACACCTTCTGTTGAGACATTTGACPTTTTCAGAGGCC 591
 QY 685 AAGAAAATACCTGTGATGGAAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGAACCTACC 744
 DB 592 AAGAGGAACCGATTACGAAATAGATCTGAATAGTTTGGAGAGGACTGCACTTGACA 651
 QY 745 ATGACACAGCTCAATGATTTGTGATCTGCTGCTGATGATGATGATGATGATGATGATG 804
 DB 652 ATAGAGCAGTTGTTGATCTTTGCAATATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCT 711
 QY 805 GGTATCGGGGGGCAACAGCTCTCAAACTTATTCGTCAACATGGGTCCATAGAAAGATC 864
 DB 712 GGTGTTGGTCCAGTCAGACGCTTAAAAATTGATAAAAGCGATGGATCCATCGAAAAATC 771

QY 865 TTGAGAAATCTT-----AATAAAGACAGATATCAAAATTCCTGAGGACTGG 909
 DB 772 GTGGAGTTTATTGAATCTGGGAGTCAACACACATTAATGGAATCCCAAGAGACTGG 831
 QY 910 CTTTACCAGAAGCTGCGAGCTTTGTTCAAGGAGCCCTTAATGTCACATTGGATA---TTCCT 966
 DB 832 CTTTACAACAAGCAAGAAATGCTGTTTCTTGACCTGAGTTATAGATGCTAAGCAATA 891
 QY 967 GAGCTAAATGCACTGCACCTGATGAGGAGGCTCTCATAGTTTCTCTGGTAAAGATAAT 1026
 DB 892 AACTTGAATGTCGCCACCAAGGAGGAGAACTTATCGAGTATTTATGTGATGATAAG 951
 QY 1027 GCTTTCAACAGATCGGTGTGACAAAGGCCATAGACAAGATCAAACTGCAAGAAATAA 1086
 DB 952 AATTCAGTGAAGAAGAGCTTAATCTGTTATATCAAGATTGAANAAGCTTGAAATCT 1011
 QY 1087 TCGTCCGAAGAGACTCGAGTCTTTTTCAGCCCACTGCCACCATCAGACCCGCT 1145
 DB 1012 GGCATTGAGGTAGTTAGATGGTCTTCCAAAGTGTGCTTAAGACAAAGAACAGCT 1070

RESULT 11

ABL14286
 ID ABL14286 standard; cDNA; 3362 BP.

XX ABL14286;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37340.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB70183.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions

XX Claim 1; SEQ ID NO 37340; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3362 BP; 979 A; 709 C; 750 G; 924 T; 0 other;

Query Match 17.3%; Score 252.6; DB 23; Length 3362;
 Best Local Similarity 55.2%; Pred. No. 1.2e-53;
 Matches 536; Conservative 0; Mismatches 429; Indels 6; Gaps 2;

QY 166 GGCGCGAAATCGCGTCGACGCGCAGCATGAGCATATACAGTTCCTGATGTAGTTGGA 225
 Db 1144 GGTGCGAAGTAGCAATCGATGCTAGTATGCTGTACCATTCCTCATCGCGTCGCG 1203

QY 226 AGGACAGGATGGAATCTCTCAAAATGAAGCTGGTGAAGTCACTAGTCTTTCGAAGGA 285
 Db 1204 TCGGAAGGC---GCCAGTTGGCCACCGTAATGGTGTATCCACGTCGCCACTTAATGGGC 1260

QY 286 ATGTTCAACCGGACATAAGATTACTTGAAGCGGAGATCAAGCCAGTTTATGTTTTCAT 345
 Db 1261 ATGTTTCAACCGGACCATCCAGATTGCTGGACACGGAATCAAGCCGCTATATGTTCTGCAC 1320

QY 346 GGCAAGCCTCTGATATGAAGAACAAAGAGCTTGTCTTAAAGATACCTCAAAAGAGATGAT 405
 Db 1321 GGAAGCCACAGATCTCAAGTCCGGTGAGCTGGCGAAGCGCGCGGAGGAA 1380

QY 406 GCAACCAAGATCTGACTGAGGCGATGAGGTAGGATGAAGATGCGATTGAAAAATG 465
 Db 1381 GCGGAGAAGGCACTGAAGCGCGCCAGCATGCGGGAGATGATGCCGTATCGAAAAAGTTT 1440

QY 466 AGCAAGAGGACTGTAAAGGTCAAGGCAACACACAGCAAGATTGTAACCGCTATTAGA 525
 Db 1441 AATCGCGGATTGTCGGGTAAAGGAGGACGACGCAAGAGGCCCAAGGAAGTCTACACA 1500

QY 526 CTTATGGGGTTCCTGTTGTAGAGGACCTTCTGAAGCAGAAAGCAAGATGTCAGCGCCTT 585
 Db 1501 CTAATGGGTGCGCTTGTGATGACCGTGCAGCGGAGCGCCAGTGTGCCGCTCTG 1560

QY 586 TGCATAACGATGAAGTGTTCGCTGTTGCTTCAAGATGATGACATGACCTTTTGGG 645
 Db 1561 GTGAAAGCTGGAAGGTTTATGCGACCGCGCACGAGGATGATGATGCTTACATTCGGA 1620

QY 646 GCTCCACGGTTCCTGCTGATTAATGATGATCAAGTTCGAAGAAATACCTGTGATGGA 705
 Db 1621 TCTACAAATGTTGATGATACCTTACAGCGAGGACGAAAGATGCGCGTCAAGAG 1680

QY 706 TTTGATTTGCCAAGTTTGGAGAGCTTTGAACCTCACCATGACGAGTTCATTGATTTG 765
 Db 1681 TTCAGCTACGACAAGCTGTTGGAAGCTCTGGCCATTAAACAATCGAGATTCATTGATCTA 1740

QY 766 TGCATCTGTGTGGATGCTGACTATTGTGATAGCATCAAGGATATCGGGGGGCAACAGCT 825
 Db 1741 TGTATTCTGCTGGGTTGCGATTACTGTGAGCATCAAGGATATTGGACCCAGCGAGCG 1800

QY 826 CTGAAACTTATTCGTCACATGGTCCATAGAAAGCATCTTGAGATCTTATAAGAC 885
 Db 1801 ATCGAATGATCAACACCTATCGGGATATAGAGACTATTCTGGATAACCTGGACTTAGC 1860

QY 886 AGATATCAAAATTCCTGAGGACTGGCTTACCAAGAAAGCTCGAGCTTGTTCAGAGAGCT 945
 Db 1861 AAATACACCGTGCAGGAACTGGAATACAGTGTGCGCGGAACTCTTCATCGAACCG 1920

QY 946 AATGTCAC---ATTGATATTCCTGAGCTTAAATGAGCTGCACCTGATGAGGAGGCTTC 1002
 Db 1921 GAGGTAGCTGATGCCGACTCCATAGATCTCAATGGGTGAGCGGATGAGGAGGCTT 1980

QY 1003 ATAAGTTTCTGTTAAAGATATGTTTCAAGGAGATCGGTGACAAAGCCATAGAG 1062
 Db 1981 GTCAGATTCTCTGCGGCGACCGGCTAGTTTCAAGGAGCGGTTTCGCAACCGTGCACAA 2040

QY 1063 AAGATCAAAATTCGCAAGAAATAAATCGTCGCAAGGAAAGCTCGAGTCTCTTTTCAAGGCA 1122
 Db 2041 AAGCTGATGAATTCGAAGCAGGCGCCAGACTCAGGTGAGACTCGATAGCTCTTTAAGACA 2100

QY 1123 ACTGCCACCAC 1133
 Db 2101 CTGCCAGCAC 2111

RESULT 12
 ABL10094
 ID ABL10094 standard; cDNA; 5471 BP.
 XX AC ABL10094;
 XX DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 24764.
 XX DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW Drosophila melanogaster.
 OS WO200171042-A2.
 XX PN 27-SEP-2001.
 XX PD 23-MAR-2001; 2001WO-US09231.
 XX PF 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX WIPI: 2001-656860/75.
 XX P-PSDB; ABB65991.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 24764; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL101840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other;

Query Match 11.8%; Score 173.2; DB 23; Length 5471;
 Best Local Similarity 54.0%; Pred. No. 1.8e-33;
 Matches 377; Conservative 0; Mismatches 318; Indels 3; Gaps 1;

QY 439 GGAGATAAGATGCGGATGAAAATTTGACAGAGAGACTGTAAAGTCAACAGGCAACAC 498
 Db 4 GGAGATGATGCGGATGCGGATGAAAAGTTAATCGCGGATTTGTCGGGTAACAGAGGACAC 63

QY 499 AACGAGATTTGAAAGCGCTATTAGACTTATGGGGTTTCTGTTGTAGAGGCACTTCT 558
 Db 64 GCCAAGAGGCGCAAGGAAGTGTCTACACTAATGGGTTGTCCTATGTTGATGACCGTGC 123

QY 559 GAAGCAGAAGCAGAAATGTGCAGCCCTTTTCATAAAGATAAGGTTCGCTGTTGCTTCA 618
 Db 124 GAAGCGAGGCGCCAGTGTCCGCTCTGTTGAAAGCTGGAAGGTTTATGCCACCGCAGC 183

QY 619 GAAGATATGACTCCCTTACTTTTGGGGCTCCAGGTTCCCTTCGTCATTTATGATFCA 678
 Db 184 GAGGATATGGATGCCCTCATTTCGGATACAAAACACTGTTGAGATACCTTACTACAGC 243

QY 679 AGTTCCAGAAATACCTGTTGATGGAATTTGTTGCCAAGGTTTGGGAGGAGCTTGAA 738

Db 244 GAGGACGAAAGATGCCCGTCAGAGTTCACCTAGGACAAAGCTGTGGAAGTCTGGCC 303
 Qy 739 CTCACATGGACCAAGTTCATTGATTTGGATCCCTGCTGTGGATGACATTTGTGATAGC 798
 Db 304 ATTAACAATCGAGAGTTCATTGATCTATGATTTCTGCTGGTTCGATTTACTGTGAGAGC 363
 Qy 799 ATCAAGGTATCGGGGGCAACACGCTCTGAACTTATTTCGTCAACATGGTCCATAGAA 858
 Db 364 ATCAAGGTATTTGGACCAAGCGGATCGAAGTTCGATCAACACCTATCGGATATAGAG 423
 Qy 859 ASCATCTTTGGAGATCTTAATAAGACAGATATCAAAATTCCTGAGGACTGGCCCTTACCAA 918
 Db 424 ACTATCTGGATAACCTGGACTCTAGCAATACACCGTCCCGGAGAACTGGAATACAAG 483
 Qy 919 GAAGTCGACGGTCTGTCAGGAGCTTAATGTCACATTTGATATTCTCT---GAGCTAAAA 975
 Db 484 GTGGCGCGGAACCTCTTCATCGAACCGGAGGTAGCTGATCGCGACTCCATAGATCTCAA 543
 Qy 976 TGGACTGCACCTGATGAGGAGGTCTCATAAGTTTCTGTAAGATATATGTTTCAAC 1035
 Db 544 TGGTTCGAGCCGATGAGGAGGCTTGTCAAGTTTCTGCGGCGACCGGCAAGTTCAAC 603
 Qy 1036 GAAGATCGGGTGACAAAGCCCATAGAGATCAAAATTCGCCAAGAAATTAATCGTCGCAA 1095
 Db 604 GAAGAGCGGTTTCGCAACGCTGCCAAAAGCTGATGAATCCAAAGCAGGCCCACTCAG 663
 Qy 1096 GGAAGACTCGATCTCTTTTCAAGCCCACTGCCACCAC 1133
 Db 664 GTGAGACTCGATGACTTCTTTAAGACACTGCCCAGCAC 701

RESULT 13

AAAT76685
 ID AAAT76685 standard; DNA; 1023 BP.

AC AAAT76685;

DT 14-APR-1998 (first entry)

XX Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.

KW Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;

KW invader directed cleavage; FEN-1; endonuclease; ds.

XX Pyrococcus furiosus.

XX WO9727214-A1.

PD 31-JUL-1997.

XX 22-JAN-1997; 97WO-US01072.

XX 02-DEC-1996; 96US-0759038.

PR 24-JAN-1996; 96US-0599491.

PR 12-JUL-1996; 96US-0682853.

PR 29-NOV-1996; 96US-0756386.

PR 02-DEC-1996; 96US-0758314.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;

PI Olive DM, Prudent JR;

XX WPI; 1997-393613/36.

DR P-PSDB; AAW24216.

XX Thermotable structure-specific nuclease(s) - used for detection and
 PT characterisation of nucleic acid sequences and variations in nucleic
 PT acid sequences

XX Example 28; Page 283-285; 457pp; English.

XX This sequence comprises the coding region of the gene encoding

CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAW24216).
 CC It was obtained by PCR amplification (see AAT76682-83). Large
 CC scale preparation of recombinant Pfu FEN-1 was performed using
 CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be
 CC used in novel methods for the detection and characterisation of
 CC nucleic acid sequences and variations in nucleic acid sequences.
 XX
 SQ Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

Query Match 10.9%; Score 159.6; DB 18; Length 1023;
 Best Local Similarity 54.6%; Pred. No. 2.5e-30;
 Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

Qy 134 CGATGAAGGACGAGAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGTCGACGCAGCA 193
 Db 29 CAAGAAAGAATAATGAGTTAGAAAACCTATACGGGAAAAAATCGAATCGAGCTCTTA 88
 Qy 194 TGACATATACCACTTCTGATTTGAGTTGGAAGGACGAGGATGAACTCTCAAAATG 253
 Db 89 ATGCAATCTACCAATTTTGTCCACAATAAGACAGAAAAGATGGAATCCACCTATGATT 148
 Qy 254 AAGCTGGTGAAGTCACTAGTCTATTTGCAAGGAATCTTCAACCGGACAAATAAGATTACTG 313
 Db 149 CAAAGGGTAGAATAACCTCCACCTTAAGCGGCTCTTTTACAGGACAATAAACCTTAATG 208
 Qy 314 AAGCGGAATCAAGCCAGTCTTATGTTTGTGCAAGCTCTCTGATATGAAGAAACAAG 373
 Db 209 AGGCTGGNATAAACCTGTGTATGTTTGTGAGAGAACTCCAGAAATCAAAAAGAAAG 268
 Qy 374 AGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCGAGTAG 433
 Db 269 AGCTCGAAAAAAG 328
 Qy 434 AGGTAGGAGATAAGATCGGATTTGAAAATTTGAGCAAGAGGAGCTGTAAAGGTCAACAAG 493
 Db 329 AAAAAGGAGATAGAGGAAGCAAGAAATATGCCCAAGAGCAACCGGTTAAATGAAA 398
 Qy 494 AACACACGAAGATTTGAAACGGCTATTAAAGCTTTATGGGGTTCCTGTGTAGAGGCAC 553
 Db 389 TGCTCATCGAGGATGCAAAAACCTTTAGAGCTTATGGGAATTCCTATAGTTCAAGCAC 448
 Qy 554 CTCTCGAAGCAGACGAGATGTGAGCCCTTTGCAATAACGATAGGTTCGCTGTTG 613
 Db 449 CTAGCGAGGAGAGGCCCAAGCTGATATATGCGCGCAAGGGGAGCGGTGATGCAATCGG 508
 Qy 614 CTTCAGAAGATATGACTCCCTTACTTTTGGGCTCCACGGTTCCCTTCGTCATTTAATGG 673
 Db 509 CTAGTCAAGATTACGATTCCTTACTTTTGGAGCTCCAAGACTTGTAGAACTTAACA 568
 Qy 674 ATCCAAGTTCCAAGAAATACCTGTGATGGAATTTGATGTTG 715
 Db 569 TAACAGGAAAAAGAAAGTTGCTCGGAAAAATGCTCTACGTCG 610

RESULT 14

AAV65840
 ID AAV65840 standard; DNA; 1023 BP.

AC AAV65840;

XX 02-FEB-1999 (first entry)

XX Pyrococcus furiosus FEN-1 endonuclease gene ORF.

XX Nucleic acid detection; multiple sequential invasive cleavage;
 KW FEN-1; endonuclease; nuclease; ds.

XX Pyrococcus furiosus.

XX WO9842873-A1.

XX 01-OCT-1998.

XX

PF 24-MAR-1998; 98WO-US05809.
XX
PR 24-MAR-1997; 97US-0823516.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
PI Mast AL, Vavra SH;
XX
XX WPI; 1998-557036/47.
DR P-PSDB; AAW79970.
DR
XX
XX
PT Detecting target nucleic acid by sequence-specific cleavage of
PT complex with two specific oligonucleotides - used to detect
PT cytomegalovirus DNA
XX
XX Example 28b; Page 316-317; 524pp; English.
XX
XX This is the nucleotide sequence of the open reading frame encoding
CC the FEN-1 endonuclease (see AAW79970) of *Pyrococcus furiosus* (Pfu).
CC It was obtained from genomic DNA by PCR amplification (see
CC also AAW5838-39). The PCR product has been ligated into vector
CC pRc99a, and FEN-1 was expressed in *E. coli* cells. The invention
CC relates to means for the detection and characterisation of nucleic
CC acid sequences, and variations in nucleic acid sequences. It also
CC relates to methods for forming a nucleic acid cleavage structure on
CC a target sequence and cleaving this structure in a site-specific
CC manner, preferably using a thermostable structure-specific nuclease
CC such as FEN-1. Cleavage of the cleavage structure by the
CC nuclease indicates the presence of specific nucleic acid sequences
CC or specific variants. The invention further relates to methods for
CC the separation of nucleic acid molecules based on charge, methods
CC for the detection of non-target cleavage products via the formation
CC of a complete and activated protein binding region, and methods for
CC the detection of nucleic acid from various viruses (e.g. human
CC cytomegalovirus) in a sample. The method amplifies the detection
CC molecule rather than the target itself, is less subject to
CC contamination than exponential amplification processes, and allows
CC many targets to be analysed in a single reaction.
XX
XX Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

Query Match 10.9%; Score 159.6; DB 19; Length 1023;
Best Local Similarity 54.6%; Pred. No. 2.5e-30;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAGGAGCAGATGTCGAGAGCTACTTCGGCGCAAAATCGCGTCGACGCCAGCA 193
DB 29 CAAGAAAGAAATTTGAGTTAGAAACCTATACGGGAAAAAATCGCAATCGACGCTCTTA 88
QY 194 TGAGCATATACCAATGCTCTGTTAGTTGGAAGCAGCAGCATGGAACCTCACAATG 253
DB 89 ATGCAATCTACCAATTTTGTCCAAATAGACAGAAAGATGGAATCCACCTATATGAT 148
QY 254 AAGCTGGTGAAGTCACTAGTCTATTTTCAAGGAATGTTCAACCGGACAATAAGATTACTGG 313
DB 149 CAAAGGTAGATAACCTCCACCTAAGCGGCTCTTTTACAGGACAATAACCTAATGG 208
QY 314 AAGCGGATCAAGCCAGTTATGTTTTGTGAGGCAAGCTCTCTGATATGAGAAACAAG 373
DB 209 AGCTGGAATAAACCCTGTGTATGTTTTGTGAGGAACCTCCAGAAATTCAAAAAGAAAG 268
QY 374 AGCTGCTCTAAAAGATACTCAAAAGAGATGTCACCAACCAAGATCTGACTGAGGCACTAG 433
DB 269 AGCTGCAAAAAGAGAGAGAGGAGAGAGAGAGAGTGAAGAAGTGGAGAGAGCACTTG 328
QY 434 AGGTAGGAGATAAAGATGCGATTTGAAAAATTTGAGCAAGAGGAGACTGTAAAGGTCAACAGGC 493
DB 329 AAAAAGGAGATAGAGGAAGCAAGAAATAATGCCCCAAGAGCAACACGGTAAATGA 388
QY 494 ACACACCAACAGATTTGAACCGCTATTAAAGACTTATGGGGTCTCTGTTAGAGGCAC 553
DB 389 TGCTCATCGAGGATGCAAAAAAAGCTCTTAGAGCTTATGGGAATTCCTATAGTTCAAGCAC 448

QY 554 CTTCTGAAGCAGACAGCAATGTGCAGCCCTTTCATATAACGATAAGTGTTCGCTGTTG 613
DB 449 CTACGGAGGAGAGGCCCAAGCTCATATATGCGCCGAAGGGAGCGTGTATCATCGG 508
QY 614 CTTCAGAGATATGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCATTAAATGG 673
DB 509 CTAGTCAAGATTACGATTCCTACTTTTGGAGCTCCAAGACTGTTTGAACAATTACAA 568
QY 674 ATCCAAGTTCAGAAATAATACCTGTGTGATGAATTTGATGTTG 715
DB 569 TAACAGAAAAAAGAAAGTTCCTGGGAAAAATGTCTACGTCG 610

RESULT 15

AAV53951
ID AAV53951 standard; DNA; 1023 BP.

XX
AC AAV53951;

XX
DT 21-DEC-1998 (first entry)

XX
DE Nucleotide sequence of the Pfu FEN-1 endonuclease.

XX Pfu FEN-1 endonuclease; Tag gene; structure-specific nuclease;
KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;
KW hepatitis C virus; HCV; thermotable; ds.

XX
OS *Pyrococcus furiosus*.

XX
FH Key Location/Qualifiers
FT CDS 1..1023

FT /tag= a
FT /product= "Pfu FEN-1 endonuclease"

XX
WO9823774-A1.

XX
PD 04-JUN-1998.

XX
PF 26-NOV-1997; 97WO-US21783.

XX
PR 02-DEC-1996; 96US-0758314.

XX
PR 29-NOV-1996; 96US-0757653.

XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX
PI Kaiser MW, Lyamichev VI, Lyamicheva N;

XX
PI WPI; 1998-322748/28.

XX
DR P-PSDB; AAW59940.

XX
PT Thermostable structure-specific nuclease(s) derived from mutant DNA

PT polymerase(s) - useful for detecting mutant allele(s) or strains of

PT microorganisms

XX
PS Example 29; Pages 278-279; 472pp; English.

XX
CC This is the nucleotide sequence encoding the Pfu FEN-1 endonuclease,
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).

XX
SQ Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

Query Match

Best Local Similarity 10.9%; Score 159.6; DB 19; Length 1023;

Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 19:50:47 ; Search time 95 Seconds
(without alignments)
4722.822 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
Sequence: 1 caccgagaatagctcgccgccc.....aaaaaaaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1463	100.0	1463	4	US-09-426-557-1
2	1421.8	97.2	1478	4	US-09-426-557-7
3	1409.8	96.4	1541	4	US-09-426-557-3
4	1350.2	92.3	1381	4	US-09-426-557-5
5	302.6	20.7	1144	2	US-08-455-968E-2
6	302.6	20.7	1144	2	US-08-455-968E-28
7	276.6	18.9	2033	2	US-08-455-968E-9
8	272.2	18.6	1930	2	US-08-455-968E-4
9	270.2	18.5	1149	2	US-08-455-968E-6
10	159.6	10.9	1023	2	US-08-757-653-175
11	159.6	10.9	1023	2	US-08-823-516-78
12	159.6	10.9	1023	3	US-08-759-038-114
13	159.6	10.9	1023	3	US-08-758-314-114
14	135.6	9.3	1032	4	US-09-146-319-1
15	135.6	9.3	1032	4	US-09-175-973-1
16	89.2	6.1	981	2	US-08-757-653-171
17	89.2	6.1	981	3	US-08-823-516-74
18	89.2	6.1	981	3	US-08-759-038-110
19	89.2	6.1	981	3	US-08-758-314-110
20	65.2	4.5	7218	1	US-08-232-463-14
21	47	3.2	455	1	US-08-636-928-4
22	47	3.2	1736	3	US-09-182-816-22
23	47	3.2	1736	3	US-09-182-816-24
24	47	3.2	1736	3	US-09-471-528-22
25	47	3.2	1736	3	US-09-471-528-24
26	47	3.2	1736	4	US-09-634-530-22
27	47	3.2	1736	4	US-09-634-530-24

```

c 28 46.8 3.2 2196 4 US-09-149-476-163 Sequence 163, Appl
 29 46.4 3.2 685 4 US-09-227-357-66 Sequence 66, Appl
 30 46 3.1 3245 3 US-09-005-180A-2 Sequence 2, Appl
 31 45.4 3.1 1359 4 US-09-387-574-11 Sequence 11, Appl
 32 45.4 3.1 1359 4 US-09-668-096-11 Sequence 11, Appl
 33 45.2 3.1 873 4 US-09-475-316A-20 Sequence 20, Appl
 34 45.2 3.1 2205 3 US-08-888-077A-41 Sequence 41, Appl
 35 45 3.1 222 4 US-08-481-190-15 Sequence 15, Appl
 36 45 3.1 222 5 PCT-US93-00869-15 Sequence 15, Appl
 37 44.4 3.0 1159 4 US-09-410-464-14 Sequence 14, Appl
 38 44.4 3.0 3410 4 US-09-020-956-110 Sequence 110, App
 39 44.4 3.0 3410 4 US-09-030-607-110 Sequence 110, App
 40 44.4 3.0 3410 4 US-09-605-785-110 Sequence 110, App
 41 44.4 3.0 3410 4 US-09-439-313-110 Sequence 110, App
 42 44.4 3.0 3410 4 US-09-352-616A-110 Sequence 110, App
 43 44.4 3.0 3410 4 US-09-602-877A-100 Sequence 100, App
 44 44.4 3.0 3410 4 US-09-232-149A-110 Sequence 110, App
 45 43.8 3.0 1223 3 US-09-154-874-4 Sequence 4, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-426-557-1
; Sequence 1, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426.557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)...(1221)
US-09-426-557-1

```

```

Query Match 100.0%; Score 1463; DB 4; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACGAGAAATAGCTCGGCGCGCGGTTTCTTGGCGCACTCCGGCTCAGCGCGCGCGGCA 60
Db 1 CACGAGAAATAGCTCGGCGCGCGGTTTCTTGGCGCACTCCGGCTCAGCGCGCGCGGCA 60
Qy 61 CCCGCCACAGCCGCGGAGAGATGGGCATCAAGGTTTGACGAACTGCTGCGGAC 120
Db 61 CCCGCCACAGCCGCGGAGAGATGGGCATCAAGGTTTGACGAACTGCTGCGGAC 120
Qy 121 AATCGGCCAAGGCGATGAAGGAGCAGAAAGTTTCGAGAGCTACTTTCGCCGCAAAATCGCC 180
Db 121 AATCGGCCAAGGCGATGAAGGAGCAGAAAGTTTCGAGAGCTACTTTCGCCGCAAAATCGCC 180
Qy 181 GTCGAGCCACATGACATATACCAGTCTCTGATTGTTAGTTGAAGGAGGAGGATGAA 240
Db 181 GTCGAGCCACATGACATATACCAGTCTCTGATTGTTAGTTGAAGGAGGAGGATGAA 240
Qy 241 ACTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTTTCGAGGAATGTTCAACCGGACA 300
Db 241 ACTCTCAAAATGAAGCTGGTGAAGTCACTAGTCACTTTCGAGGAATGTTCAACCGGACA 300
Qy 301 ATAGATTACTGGAGCGGGAATCAAGCCAGTTATGTTTGTATGCAATCCCTCTGAT 360

```



```

QY 607 GCTGTTGCTTCCAGAGATGACGCCCTTACTTTTGGGGCTCCAGGTTCTCTGTCAT 666
Db 619 GCTGTTGCTTCCAGAGATGACGCCCTTACTTTTGGGGCTCCAGGTTCTCTGTCAT 678
QY 667 TTAATGGATCCCAAGTTCCTCAAGAAATACCTGTCATGGAATTTGATGTTGCCAAGGTTTTG 726
Db 679 TTAATGGATCCCAAGTTCCTCAAGAAATACCTGTCATGGAATTTGATGTTGCCAAGGTTTTG 738
QY 727 GAGGAGCTTGAATCACCATGACAGCTTCATTTGATGTTGTCATCTCTGTCGATGTGAC 786
Db 739 GAGGAGCTTGAATCACCATGACAGCTTCATTTGATGTTGTCATCTCTGTCGATGTGAC 798
QY 787 TATTGTGATAGCATCAAGGATCGGGGGCAACAGCTCTCAAACTTATTGCTCAACAT 846
Db 799 TATTGTGATAGCATCAAGGATCGGGGGCAACAGCTCTCAAACTTATTGCTCAACAT 858
QY 847 GGTCTCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGATATCAAAATTCCTGAGGAC 906
Db 859 GGTCTCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGATATCAAAATTCCTGAGGAC 918
QY 907 TGGCCTTACCAAGAGCTCGAGCTTGTTCAGGAGCCCTAATGTCTACATTTGGATTTCT 966
Db 919 TGGCCTTACCAAGAGCTCGAGCTTGTTCAGGAGCCCTAATGTCTACATTTGGATTTCT 978
QY 967 GAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAGTTTCTGTTAAAGATAAT 1026
Db 979 GAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAGTTTCTGTTAAAGATAAT 1038
QY 1027 GGTTCACAGAGATCGGTGTACAAAGGCCATAGAGAATCAAAATCTGCCAAGATAAA 1086
Db 1039 GGTTCACAGAGATCGGTGTACAAAGGCCATAGAGAATCAAAATCTGCCAAGATAAA 1098
QY 1087 TCGTCCAGGAAGACTCGAGTCTTTTCAAGCCAACTGCCACCATCAGCACCGCTA 1146
Db 1099 TCGTCCAGGAAGACTCGAGTCTTTTCAAGCCAACTGCCACCATCAGCACCGCTA 1158
QY 1147 AAACGGAAGAGACTTCGGATAAACAAGCAAGCGAGCTCGCAACAAGAAACAAAGGCT 1206
Db 1159 AAACGGAAGAGACTTCGGATAAACAAGCAAGCGAGCTCGCAACAAGAAACAAAGGCT 1218
QY 1207 GGTGGAAGAAAGAAATTAATCTGGATGCTGTGATGATGATGATGATGATGATGATGAT 1266
Db 1219 GGTGGAAGAAAGAAATTAATCTGGATGCTGTGATGATGATGATGATGATGATGATGAT 1278
QY 1267 TGGCGTATCATTGCTGTAGATTTTAACTCCCTGTTTAACTCAGAGCTTTGGTAAA 1326
Db 1279 TGGCATGATCATTGCTGTAGATTTTAACTCCCTGTTTAACTCAGAGCTTTGGTAAA 1338
QY 1327 AGTTTGTCTATGTTTCAAGCTGGGTAAGTTAGTTGTTTGAAGAGATTGGTGTACCAA 1386
Db 1339 AGTTTGTCTATGTTTCAAGCTGGGTAAGTTAGTTGTTTGAAGAGATTGGTGTACCAA 1398
QY 1387 GTAACAAACTTATCGTGTGTTTTTACTCTTGTCTTTGAAAGTAAAGAAAGAAAGAA 1446
Db 1399 GTAACAAACTTATCGTGTGTTTTTACTCTTGTCTTTGAAAGTAAAGTAAAGTAA 1458
QY 1447 AAAAAAAAAAAAAA 1463
Db 1459 AAAAAAAAAAAAAA 1475

```

```

RESULT 3
US-09-426-557-3
; Sequence 3, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426, 557
; CURRENT FILING DATE: 1999-10-22

```

```

; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)...(1215)
; US-09-426-557-3

```

```

Query Match 96.4%; Score 1409.8; DB 4; Length 1541;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CGCGCGTTTCTTGGCCACTCCGGCTCAGCGCGCGCGCGCCGACCGCCGACGCGCGGCA 78
Db 13 CGCGCGTTTCTTGGCCACTCCGGCTCAGCGCGCGCGCGCCGACCGCCGACGCGCGGCA 72
QY 79 GAGGAGATGGCATCAAGGGTTTGACGAACTGCTGGCGGACAAATCGCCCAAGGGATG 138
Db 73 GAGGAGATGGCATCAAGGGTTTGACGAACTGCTGGCGGACAAATCGCCCAAGGGATG 132
QY 139 AAGGAGCAGAAGTTCGAGAGCTACTTTCGGCGCAAAATCGCCGCGACGCGCCAGCATGAGC 198
Db 133 AAGGAGCAGAAGTTCGAGAGCTACTTTCGGCGCGCAAAATCGCCGCGACGCGCATGAGC 192
QY 199 ATATACAGTTCCTGATTTAGTTGGNAGCAGGATGGAACCTCTCACAATATGAGCT 258
Db 193 ATATACAGTTCCTGATTTAGTTGGNAGCAGGATGGAACCTCTCACAATATGAGCT 252
QY 259 GGTGAAGTCTACTAGTTCATTTGCAAGGAATGTTCAACCGGACAAATAGATTACTCGAAGCG 318
Db 253 GGTGAAGTCTACTAGTTCATTTGCAAGGAATGTTCAACCGGACAAATAGATTACTCGAAGCG 312
QY 319 GGAATCAAGCCAGTTTATGTTTGTGATGCGAAGCCTCCTGATATGAAGAACAAGAGCTT 378
Db 313 GGAATCAAGCCAGTTTATGTTTGTGATGCGAAGCCTCCTGATATGAAGAACAAGAGCTT 372
QY 379 GCTTAAAGATCTCAAAAGAGATGATGCAACCAAGATCTGCTGAGGAGGATGAGGTA 438
Db 373 GCTTAAAGATCTCAAAAGAGATGATGCAACCAAGATCTGCTGAGGAGGATGAGGTA 432
QY 439 GGAGATAAGATGCGATTGAAAAATTCAGCAAGAGAGCTGTAAAGGTCACAAGTCAACAC 498
Db 433 GGAGATAAGATGCGATTGAAAAATTCAGCAAGAGAGCTGTAAAGGTCACAAGTCAACAC 492
QY 499 AAGCAAGATTGTAACCGGCTATTAAAGACTTATGGGGGTTCTGTTGTAGAGGACCTTCT 558
Db 493 AAGCAAGATTGTAACCGGCTATTAAAGACTTATGGGGGTTCTGTTGTAGAGGACCTTCT 552
QY 559 GAAGCAGAACGAGATCTGAGCGCTTTGATATAACCAATAGAGTGTTCGGCTGTTCTTCA 618
Db 553 GAAGCAGAACGAGATCTGAGCGCTTTGATATAACCAATAGAGTGTTCGGCTGTTCTTCA 612
QY 619 GAAGATATGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGATTAATGATGATCA 678
Db 613 GAAGATATGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGATTAATGATGATCA 672
QY 679 AGTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 738
Db 673 AGTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 732
QY 739 CTCACCATGGACAGTTCATTTGATGTTGTCATCTGTTGATGATGATGATGATGATGATG 798
Db 733 CTCACCATGGACAGTTCATTTGATGTTGTCATCTGTTGATGATGATGATGATGATGATG 792
QY 799 ATCAAGGATATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCCATGAA 858
Db 793 ATCAAGGATATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCCATGAA 852

```

Qy	859	AGCACTCTGGAGAACTTTAATAAAGACAGATATCAAAATTCCTGAGACTGGCCTTACCAA	918
Db	853	AGCATCTTGGAGAACTTTAATAAAGACAGATATCAAAATTCCTGAGACTGGCCTTACCAA	912
Qy	919	GAAGCTCGAGCTGTGTCAAGGAGCCCTAATGTCACATGGGATATTCCTGAGCTAAATGG	978
Db	913	GAAGCTCGAGCTGTGTCAAGGAGCCCTAATGTCACATGGGATATTCCTGAGCTAAATGG	972
Qy	979	ACTGCACCTGATGAGGAGGCTCTCAATAAGTTCCTCGGTAAAGAGATAATGGTTCAACGAA	1038
Db	973	ACTGCACCTGATGAGGAGGCTCTCAATAAGTTCCTCGGTAAAGAGATAATGGTTCAACGAA	1032
Qy	1039	GATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAATAAATCTGTCGAAGGA	1098
Db	1033	GATCGGTGAGAAGGCCATAGAGAAGATCAAAATCTGCCAAGAATAAATCTGTCGAAGGA	1092
Qy	1099	AGACTCGAGTCCCTTTTTTCAAGCCCACTGCCACCACATCAGCACCGCTAAACCGGAAGGAG	1158
Db	1093	AGACTCGAGTCCCTTTTTTCAAGCCCACTGCCACCACATCAGCACCGCTAAACCGGAAGGAG	1152
Qy	1159	ACTTCGGATAAAAAACGAAGGCAGCTGCGAACAAGAAAAACAAAGGCTTGGTGGAAAGAAG	1218
Db	1153	ACTTCGGATAAAAAACGAAGGCAGCTGCGAACAAGAAAAACAAAGGCTTGGTGGAAAGAAG	1212
Qy	1219	AAATAATCTTGGATGCTTGATGTACAACTACGACGTACGAAGACAGCGGTGGCGGTATCAC	1278
Db	1213	AAATAATCTTGGATGCTTGATGTACAACTACGACGTACGAAGACAGCGGTGGCGGTATCAC	1272
Qy	1279	TTGCGTTAGATATATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAAGTTTGCCTCATG	1338
Db	1273	TTGCGTTAGATATATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAAGTTTGCCTCATG	1332
Qy	1339	TTTCAAGCTGGGTAAAGTTAGTTGTGTTTTGAAGAGATTGGTGTACCAAGTAACAAAACTT	1398
Db	1333	TTTCAAGCTGGGTAAAGTTAGTTGTGTTTTGAAGAGATTGGTGTACCAAGTAACAAAACTT	1392
Qy	1399	ATCGCTGTTTTTTTACTCTCTTGTGCTTTGAAGTA	1431
Db	1393	ATCGCTGTTTTTTTACTCTCTTGTGCTTTGAAGTA	1425

RESULT 4

```

RES001 4
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 623257
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5

```

Query Match 92.3%; Score 1350.2; DB 4; Length 1381;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 8; Indels 0;

[illegible]

QY	109	CTGCTGGCGGACAATGCGCCCAAGCGGATGAAGGAGCAGAAAGTTCGAGAGCTACTTCTCGCG	166
DB	61	CTGCTTGGCGGACAAATGCGCCCAAGCGGATGAAGGAGCAGAAAGTTCGAGAGCTACTTCTCGCG	120
QY	169	CGCAAAATCGCGCTCGAGCCGAGCATGAGCATATACCAGTTCTCGTGAATGTAGTTGGAAGG	228
DB	121	CGCAAAATCGCGCTCGAGCCGAGCATGAGCATATACCAGTTCTCGTGAATGTAGTTGGAAGG	180
QY	229	ACAGCATGGAAACTCTCACAAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATG	288
DB	181	ACAGCATGGAAACTCTCACAAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATG	240
QY	289	TTCAACCGGCAAAATAAGATTACTCGAAGCGGGATCAAGCCAGTTTATGTCTTTTGAATGCG	348
DB	241	TTCAACCGGCAAAATAAGATTACTCGAAGCGGGATCAAGCCAGTTTATGTCTTTTGAATGCG	300
QY	349	AAGCCTCCTGATATGAAGAAACAAGACTTTGCTAAAAGTACTCAAAAAGAGATGATGCA	408
DB	301	AAGCCTCCTGATATGAAGAAACAAGACTTTGCTAAAAGTACTCAAAAAGAGATGATGCA	360
QY	409	ACCAAGATCTGACTGAGGCACTAGAGGTAGGAGATAAAGATCGGATTCGAAAAATTGAGC	468
DB	361	ACCAAGATCTGACTGAGGCACTAGAGGTAGGAGATAAAGATCGGATTCGAAAAATTGAGC	420
QY	469	AAGAGGACTGTAAAGGTTCACAAGGCAACAACGAAAGATTGTAAACGGCTATTAAGACTT	528
DB	421	AAGAGGACTGTAAAGGTTCACAAGGCAACAACGAAAGATTGTAAACGGCTATTAAGACTT	480
QY	529	ATGGGGGTTCTGTTGTAGAGGCACTTTCTGAAGCAGAAAGAGATGTGCGACCCCTTTGC	588
DB	481	ATGGGGGTTCTGTTGTAGAGGCACTTTCTGAAGCAGAAAGAGATGTGCGACCCCTTTGC	540
QY	589	ATAAACGATAAGGTGTTGCGCTGTGCTTCAGAAAGATATGCACTCCCTTACTTTTGGGCT	648
DB	541	ATAAACGATAAGGTGTTGCGCTGTGCTTCAGAAAGATATGCACTCCCTTACTTTTGGGCT	600
QY	649	CAOAGGTTCTCTCTCATTTAATGGATCCAAGTTCGAAGAAATACCTGTGATGGAATTT	708
DB	601	CAOAGGTTCTCTCTCATTTAATGGATCCAAGTTCGAAGAAATACCTGTGATGGAATTT	660
QY	709	GATGTTGCCAAGGTTTGGAGGAGCTTGAACTCACCATGGACAGTTCATGATTTGTGC	768
DB	661	GATGTTGCCAAGGTTTGGAGGAGCTTGAACTCACCATGGACAGTTCATGATTTGTGC	720
QY	769	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGGCAACAGACTCTG	828
DB	721	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGGCAACAGACTCTG	780
QY	829	AAACTTATTGCTCAACATGGGTCCATAGAAAGCATCTTGAGAAATCTTAATAAGACAGA	888
DB	781	AAACTTATTGCTCAACATGGGTCCATAGAAAGCATCTTGAGAAATCTTAATAAGACAGA	840
QY	889	TATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAAGCTTCGACGCTTGTTCAGGAGGCTAAT	948
DB	841	TATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAAGCTTCGACGCTTGTTCAGGAGGCTAAT	900
QY	949	GTCACATGGGATATTCCTGAGCTAAATGGAGCTGCACCTGTATGAGAGGGTCTCATAGT	1008
DB	901	GTCACATGGGATATTCCTGAGCTAAATGGAGCTGCACCTGTATGAGAGGGTCTCATAGT	960
QY	1009	TTCTCTGTAAAGATATGTTTCAACGAAGATCGGGTTCACAAAGCCATAGAGAAATC	1068
DB	961	TTCTCTGTAAAGATATGTTTCAACGAAGATCGGGTTCACAAAGCCATAGAGAAATC	1020
QY	1069	AAATCTGCCAAGAAATAATCTGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCCAACTGCC	1128
DB	1021	AAATCTGCCAAGAAATAATCTGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCCAACTGCC	1080
QY	1129	ACCACATCAGCCGCTTAAACGGAGGAGACTTCGGATATAACAAAGCAAGGAGCTGCG	1188
DB	1081	ACCACATCAGCCGCTTAAACGGAGGAGACTTCGGATATAACAAAGCAAGGAGCTGCG	1140
QY	1189	AACAAGAAAAAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAACATA	1248

Db 1141 AACAGAAACAAAGGCTGTGGAAAGAAATAATCTTGGATGCTTGTATACAACTA 1200
QY 1249 CGACTAGAAAGACGCGGTGGCTGATCAGCTTGCCTTAGATTAATTAACCTCCCTGTTTAA 1308
Db 1201 CGACTAGAAAGACGCGGTGGCTGATCAGCTTGCCTTAGATTAATTAACCTCCCTGTTTAA 1260
QY 1309 ACTCAGAGCTTTGGTAAAGCTTGTCTCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG 1368
Db 1261 ACTCAGAGCTTTGGTAAAGCTTGTCCCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG 1320
QY 1369 AAGAGATTGGTACCAGTAACAAAACCTTATCGCTGTTTTT 1411
Db 1321 AAGAGATTGGTACCAGTAACAAAACCTTATCGCTGTTTTT 1363

RESULT 5

US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-455-968E-2

Query Match 20.7%; Score 302.6; DB 2; Length 1144;
Best Local Similarity 56.6%; Pred. No. 7.5e-70;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 85 ATGGGCATCAAGGGTTTGACGAACACTGCTGGCGGACAAATCGCCCAAGCGCATGAAGGAG 144
Db 1 ATGGGAATTCGAAGCCTGGCCCAACTAATGCTGATGGCCCGCCAGTGCCATCCGGGAG 60
QY 145 CAGAAGTTGAGAGCTACTTCGGCCGCAAAATCCCGCTGCGACGACAGCATGATATAC 204
Db 61 AATGACATCAAGAGCTACTTTGGCCGTAAGGTGGCCATTTGATGCTCTATGAGCATTTAT 120
QY 205 CAGTTCTGATTGATGTTGGNAGACAGGCAATGGAACCTCTCACAATGAAGCTGGTGA 264
Db 121 CAGTTCTGATTGTTGTCGCCAG---GGTGGGATGTGCTGCAGAAATGAGAGGGTGAG 177

QY 265 GTCACTAGTCATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTGGAGCGGGAATC 324
Db 178 ACCACCAAGCACCTGATGGGCATGTTCTACCGCACCATTCCATGATGGAGAACGGCATC 237
QY 325 AAGCCAGTTTATGTTTTTGTATGGCAAGCCTCCTGATATGAAGAACAAAGAGCTTGTCTAAA 384
Db 238 AAGCCCGTGTATGTTTGTATGGCAAGCCTCCTGATATGAAGAACAAAGAGCTTGTCTAAA 297
QY 385 AGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGAGTAGAGGTAGGAGAT 444
Db 298 CGCAGTGAAGCGCGGCTGAGGCAGAGAAGCAGCTGCAGCAGGCTCAGGCTGCTGGGGCC 357
QY 445 AAGATCGCATTTGAAAAATTTGAGCAAGAGGACTCTAAAGGTCAACAAGGCAACAACAAGCA 504
Db 358 GAGCAGAGGTGGAAAAATTTCACTAAGCGGCTGTGAGGTCACCTAAGCACCACATGAT 417
QY 505 GATTGTAACGGCTATTAAAGACTTTATGGGGTTCCTGTTGAGAGGACCTTCTTGAAGCA 564
Db 418 GAGTCAAAACATCTGCTGAGCCTCATGGCATCCCTTATCTTTGATGACCCAGTGAGGCA 477
QY 565 GAAGCAGAATGTGCAGCCCTTTGTCATAAAGATTAAGGTGTCGCTGTTGCTTCAGAGAT 624
Db 478 GAGGCCAGCTGTGCTGCCCTGGTGAAGGTGGCAAGTCTATGCTGGGCTACCGAGGAC 537
QY 625 ATGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTGTCTATTTAATGGATCCAAGTTCC 684
Db 538 ATGACTGCCCTACCTTCGGCAGCCCTGCTCTAATGCGACACCTGACTGC-AGTGAAGCC 597
QY 685 AAGAAAATACCTGTGATGGAAATTTGATGTTGCCAAGGTTTGGAGGAGCTTGAACCTACC 744
Db 598 AAAAGCTGCCAATCCAGGAATTCACCTGAGCGCGGATTCCTGCAGGAGCTGGGCTGAAC 657
QY 745 ATGACCAGTTCATTGATTTGTGATCTGCTGTGATGCTGACTATTTGATAGCATCAA 804
Db 658 CAGGAACAGTTTGTGGATCTGTGATCTCTCTAGGAGTACTACTGTGAGAGTATCCGG 717
QY 805 GGTATCGGGGGGCAACAGACTCTGAAACTTATTGCTCAACATGGGTCCATAGAAAGCATC 864
Db 718 GGTATCGGGGGCAACAGCGGCTGTGGACCTCATCCAGAAGCACAGAGCATCGAGAGATC 777
QY 865 TTGGAGAACTTTAATAAGACAGATATCAAAATTCCTGAGGAGCTGGCTTACCAAGAGCT 924
Db 778 GTGCGGGGAGTTGACCCCAACAAGTACCTGTGCGAGAAAATTTGGCTCCACAAGGAGCT 837
QY 925 CGAGCTTTGTTCAAGGAGCCCTAATG---TCACATTTGGATATTCTGAGCTAAAATGGACT 981
Db 838 CACCAGCTCTTCTTGGAACTTGAGTGTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 897
QY 982 GCACCTGATGAGGAGGTCTCATAAAGTTTCTGTTAAAGATAATGTTTTCACGGAAGAT 1041
Db 898 GAGCCAAATGAAGAAGAGCTGATCAAGTTTCATGTGTGGTGAAGAGCAGTTCTCTGAGGAG 957
QY 1042 CGGTGACAAGGCCCATAGAGAATCAAAATCTGCCAAGATAAATCGTCCCAAGGAAGA 1101
Db 958 CGAATCCGAGTGGGGTCAAGAGGCTGAGTAAGAGCGCCCAAGGACAGCACCCAGGGCGC 1017
QY 1102 CTCGAGTCTCTTTTCAAGCCCAACTGCCACCATCATCAGCAC 1142
Db 1018 CTGATGATTTCTTCAAGGTGACCGGCTCACTCTCTTCAGC 1058

RESULT 6

US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (polynucleotide)
US-08-455-968E-28

Query Match 20.7%; Score 302.6; DB 2; Length 1144;

Best Local Similarity 56.6%; Pred. No. 7.5e-70;

Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 85 ATGGGATCAAGGTTTGGAGAACTGCTGGGACAACTGCGCCCAAGGGGATGAAGAG 144
DB 1 ATGGGAATTCAGGCCCTGGCCAACTAATGCTGATGTGGCCCCCAGTGCCATCCGGAG 60
QY 145 CAGAAGTTCGAGAGCTACTTGGCGGCAAAATCGCCGTCAGCCGAGCATGAGCATATAC 204
DB 61 AATGACATCAAGAGCTACTTGGCGGTAAGGTGCCATGATGCTCTATGAGCATTTAT 120
QY 205 CAGTTCCTGATTTAGTTGGAAGGACAGGATGGAAGTCTCACAATGAGTGGTGA 264
DB 121 CAGTTCCTGATTTGCTGTGCCAG--GGTGGGATGCTGTCAGAAATGAGGAGGTGAG 177
QY 265 GTCACTAGTCAATTTGCAAGAAATTTCAACCGGACAAATAGATTACTGGAAGCGGAATC 324
DB 178 ACCACAGCCACCTGATGGGATGTTCTACGCACCAATTCGATGATGGAGACGGCATC 237
QY 325 AAGCCAGTTTATGTTTGTGGAAGCCCTCCTGATATGAAGAACAGAGCTTGTCTAAA 384
DB 238 AAGCCCGTGTATGCTTTGATGCAAGCCGACAGCTCAAGTCAGGCGAGCTGGCCAAA 297
QY 385 AGATACATAAAGAGATGATGCAACCAAGATGCTACTGAGGAGTGTAGGATGAGAT 444
DB 298 CGCAGTGAAGGCGGGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357
QY 445 AAGATCGGATTTGAAATTTGAGCAAGAGGACTGTAAGGTGCAAGGCAACACAAAGAA 504
DB 358 GAGCAGAGGTGGAAAATTTCACTAAGCGGCTGGTGAAGGTCACTAAGCAGCAATGAT 417
QY 505 GATTGTAAGCGCTATTAAAGCTTTATGGGGTTCCTGTTGATGAGGACACCTTCTGAAGCA 564
DB 418 GAGTGAACACATCTGCTGAGCCTCATGGCATCCCTTATCTTATGATGACCCAGTGAGGCA 477
QY 565 GAAGCAGAAATGTCAGCCCTTTGCAATAAAGATGATGATGATGATGATGATGATGAT 624
DB 478 GAGGCCAGCTGTGCTGCCCTGGTGAAGGCTGGCAAGGCTATGCTGCGGCTACCGAGGAC 537
QY 625 ATGGACTCCCTTACTTTTGGGGTCCACGGTTCCTTCCTCATTTAATGATCAAGTTCC 684

DB 538 ATGGACTGCCTCACCCTCGGAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCC 597
QY 685 AAGAAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGAACCTACC 744
DB 598 AAAAAGCTGCCAATCCAGGAATTCACCTGAGCGGATTCGACAGGAGCTGGGCTGAAC 657
QY 745 ATGGACCAAGTTCAATGATTTGTGTCATCTGCTGTGATGATGATGATGATGATGATGAT 804
DB 658 CAGGAACAGTTTGTGGATCTGTGATCTGCTAGGCACTGACTCTGTGAGAGTATCCGG 717
QY 805 GGTATCGGGGGCAACAGCTCTCAAACTTATTCGTCAACATGGTCCATGAAGAGCATC 864
DB 718 GGTATTGGCCCAAGCGGCTGTGGACCTCATCAGAAAGCAAGAGCATCGAGGATC 777
QY 865 TTGGAGAAATCTTAATAAAGCAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAAGCT 924
DB 778 GTGCGGCACTTGACCCCAACAACTACCTGTGCCAGAAAATTCGCTCCACAGGAGCT 837
QY 925 CGAGCTTGTTCAGGAGCCTAATG---TCACATTTGATATTCCTGAGCTTAAATGGACT 981
DB 838 CACCAGCTCTTCTTGGAACTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGC 897
QY 982 GCACCTGATGAGGAGGCTCAATAAGTTTCTGTTAAAGATATGTTTCAACGAAGAT 1041
DB 898 GAGCCAAATGAAGAGAGCTGATCAAGTTTATGTTGTTGAAAAGAGCTTCTCTGAGGAG 957
QY 1042 CGGGTGACAAAGGCTATAGAGAGATCAAAATCTGCCAAGATAAATCGTCCAAAGGAAGA 1101
DB 958 CGAATCCGCACTGGGTCAAGAGGCTGAGTAAGAGCCGCCAAGCAGCACCAGCGGCCG 1017
QY 1102 CTCGAGTCTTTTCAAGCCAACTGCCACCACTACAGCCTTAAATGGACT 1142
DB 1018 CTGGATGATTTCTTCAAGGTGACCGGCTCACTCTCTTTCAGC 1058
RESULT 7
US-08-455-968E-9
; Sequence 9, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

Fri May 9 16:32:33 2003

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 104...1237
US-08-455-968E-9

Query Match 18.9%; Score 276.6; DB 2; Length 2033;

Best Local Similarity 55.1%; Pred. No. 6.2e-63;
Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGATCAAGGTTTACAGAACTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGAG 144
DB 104 ATGAAATTCAGGGCTTCCCAAACTAATTTGCTGATGGGCCCGCAGTGCATCCGTGAG 163
QY 145 CAGAGTTCGAGAGTACTTTCGAGCAAAATCGCGCTGCGAGCCAGCATGAGCATATAC 204
DB 164 AATGACATCAAGAGTACTTTGGTGGCAAGTGGCCATCGATGCTCCATGAGGCTATC 223
QY 205 CAGTCTCTGATTTAGTGGAGGACAGGATGGAACCTCTCAAAATGAAGCTGTGAA 264
DB 224 CAGTCTCTGATTTAGTGGAGGACAGGATGGAACCTCTCAAAATGAAGCTGTGAA 280
QY 265 GTCACTAGTCAATTCAGCAAACTTCAACGGCAATTAAGATTACTGGAAGCGGGAATC 324
DB 281 ACCACGAGCC---TGATGGGATGTTCTACCGTACCATGCG---CATGGAAATGGCATC 334
QY 325 AAGCCAGTTTATGTTTATGATGGCAAGCTCTGATATGAGAAACAGAGCTTGCTAAA 384
DB 335 AAGCTGTGATGCTTTGATGGCAAGCTCTGATATGAGAAACAGAGCTTGCTAAA 394
QY 385 AGATACCTCAAAAGAGATGATCAACCAAGATCTGACTGAGGAGTAGAGTAGAGAT 444
DB 395 CGCAGTGAAGGGCGCCGAGGCTGAGAGCAACTGACAGGCTCAGCAGGCTGGGATG 454
QY 445 AAGATGCGATTCAGAAATTTAGCAGAGGACTGTAAGGCTCAACAGGCAACAGCAAG 504
DB 455 GAGGAGAGGTTGAGAGGTTTCAACAGAGGCTGTAAGGCTCAACAGGCAACAGCAAG 514
QY 505 GATTGTAAGCGCTTATTAAGCTTATGAGGCTTCTGTTAGAGGACCTTCTGAAGCA 564
DB 515 GAGTGAACAAACCTCTGAGGCTCATGGGCTTCTGTTAGAGGACCTTCTGAAGCA 574
QY 565 GAGCAGATGTCAGCCCTTTGCATTAACAGATAGGCTGCTGTTGTTGCTCAGAGAT 624
DB 575 GAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634
QY 625 ATGAGCTCCCTTACTTTTGGGCTCCAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 684
DB 635 ATGAGCTCCCTTACTTTTGGGCTCCAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 694
QY 685 AAGAAATACCTGATGGAATTTGATGTTGCAAGGTTTGGAGGAGCTTGAACCTACC 744
DB 695 AAGAGCTGCCATCCAGAGTTTCACTGAGCCCGCTGCTGAGGAGCTGGGCTGAAC 754
QY 745 ATGAGCAGTCAATGATTTGTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 804
DB 755 CAGGAGCAGTTTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
QY 805 GGTATCGGGGGCAACACCTGAACTTATTCCTACATGCTGCTGCTGCTGCTGCTGCTG 864
DB 815 GGCATTTGGCCCAAGCGGCTGTGATCTCATCCAGAAACATAGAGCTCGAGAGATC 874
QY 865 TTGAGATCTTAAATAAGACAGATATCAAAATTCCTGAGGAGCTGGCCTTACCAAGAGCT 924
DB 875 GTGAGCGGCTGGACCCCAAGTACCCGCTTCCAGAGAACTGGCTCCCAAGAGGCC 934
QY 925 CGAGCTTGTTCAGAGGAGCTTAA---TGTCAATTTGGATTTCTGAGCTTAAATGAGCT 981
DB 935 CAGAGCTCTTCTTGAGGACAGAGTAGTGGACCAAGAGTCTGTGGAGCTGAAGTGAGC 994
QY 982 GCACCTGATGAGAGGCTCTCAATAGTTTCTCTGTTGTTAAAGATAATGTTTCAACGAAGAT 1041

DB 995 GAGCAAAATGAAGAGAGTTGGTCAAAATTTATGTGTGGTGAAGAGAGTTTGAAGAG 1054
QY 1042 CGGTGCAAGAGGCGCATAGAGAGATCAAAATCTGCCAAGAAATAAATCGTCGCA/GAAGA 1101
DB 1055 CGAATTCGAGTGGGTCAAGCGCTGAGTAAGAGCGCCAGGCGCAGCAGCGAGC 1114
QY 1102 CTCGAGTCTTTTCAAGCCAACTGCCACCATGAGCCGCTAAACCGAGAGAGCT 1161
DB 1115 CTCGATGATTTCTTCAAGGTGACAGGCTCTACTCTCTCTCAGC-----TAAGCGC 1162
QY 1162 TCGATAAACAAAGCAGGAGCTGCGAACCAAGAAACAAAGAGCTGCTGAAGAAAGAAA 1221
DB 1163 AAGGAGCAGAACCAAGGCGCTGCTGAAGAAAGCAAGAGCTGGGGGAGGAG 1222
QY 1222 T 1222
DB 1223 T 1223

RESULT 8
US-08-455-968E-4
; Sequence 4, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1930 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-4

Query Match 18.6%; Score 272.2; DB 2; Length 1930;
Best Local Similarity 55.1%; Pred. No. 8.7e-62;
Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 85 ATGGCATCAAGGTTTGAAGAACTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGAG 144
DB 1 ATGGGAATTCAGGCTTGGCAAACTAATTTGCTGATGTCGCCCGCCATGCCATCCGTGAG 60
QY 145 CAGAAGTTTCGAGAGCTACTTTCGCGCGGCAAAATCGCGTGCAGCCAGCATGAGCATATAC 204
DB 61 AATGACATCAAGAGCTACTTTGGTTCGTAAGAGTGGCCATCGATCGCTCCATGAGCATCTAC 120

QY 505 GATTGTAACGGCTATTAGACTTATGGGGTTCCTGTTGTAGAGGACCTCTCTGAAGCA 564
DB 412 GAAGCCAAATAATTACTAGGACTAATGGGATCCCATATATATAGCGCAACGGAGCT 471
QY 565 GAAGCAGAAATGTCAGCCCTTTTCATAAACGATAAGGTGTTCCTGTGCTTCAGAAAT 624
DB 472 GAGGCTCAATGTCTGAGTTGGCAAGAGGAAAGGTGTATGCCGACCAAGTGAAGAT 531
QY 625 ATGACTCCCTTACTTTTGGGGTTCACAGGTTCCCTTCGTCTCATTTATGGATCAAGTTC 684
DB 532 ATGACACACTCTGTATAGAACCCCTTCTGTTGTGAGACATTTGACTTTTCAGAGGCC 591
QY 685 AAGAAATACCTCTGATGGAATTTGATGTTGCCAAGTTTGGAGAGCTTGAACCAACC 744
DB 592 AAGAGGACCGAATTCACGAAATAGACTGAATTTAGAGGACTCGACTTGACA 651
QY 745 ATGACACAGTCAATGATTTGTGATCCTGCTGTGATGAGTGTGATGATGATGATCAAA 804
DB 652 ATAGACAGTTTGTGATCTTTCATATATGCTTTGGTGTGACTACTGTGAAAGCATCAGA 711
QY 805 GGTATCGGGGGCAACAGACTCTGAACITTTTCGTCAACATGGTCCATAGAAAGCATC 864
DB 712 GGTGTTGGTCCAGTGACAGCCTTAAATTTGATAAAACGATGGATCCATCGAAATAATC 771
QY 865 TTGGAGAATCTT-----ATAAAGACAGATATCAAAATTCCTCGAGGACTGG 909
DB 772 GTGAGTTTATTAATCTGGGAGTCAACAACTAAATGGAAATCCAGAGACTGG 831
QY 910 CTTTACCAAGAGCTCGACGCTTTGTTCAAGGAGCCTAATGTCACATTTGGATA---TTCC 966
DB 832 CTTTACCAAGAGCTCGACGCTTTGTTCAAGGAGCCTAATGTCACATTTGGATA---TTCC 966
QY 967 GACCTAAATGAGTCACTGATGAGGAGGCTCTCAATAGTTTCTGGTAAAGATAAT 1026
DB 892 AACTTGAATGGTCCACCAAGAGGAGGAACTTATCGAGTATTTATGATGATAAG 951
QY 1027 GGTTCACCAAGAGCTGGTGACAAAGGCGCATAGAGAGATCAAACTGTCACCAAGATAA 1086
DB 952 AATTCAGTCAAGAGAGTAAATCTGGTATATCAAGTTGAAAGAGCTTGAATCT 1011
QY 1087 TCGTCCAGAGAGACTCGAGTCTTTTCAAGCCCACTGCCACCATCAGCACCGCT 1145
DB 1012 GGCATTTCAGGTTAGATGAGTCTTCCAAAGTGTGCTTAAGACAAAGGACAGCT 1070

RESULT 10

US-08-757-653-175
; Sequence 175, Application US/08757653
; Patent No. 5843669

GENERAL INFORMATION:

; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 175:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1023 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1020

; US-08-757-653-175

Query Match 10.9%; Score 159.6; DB 2; Length 1023;

Best Local Similarity 54.6%; Pred. No. 2e-32;

Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAAGGAGCAGAGTTCGAGAGCTTCTCGAGAGCTTTCGGCCGCAAAATCGCGTCCACGCCAGCA 193

DB 29 CAAGAAAGAAATTTGAGTTAGAAACCTATACGGGAAAAAATCGCAATCGACCTCTTA 88

QY 194 TGACATATACCACTTCTGATTTGAGTGTGAGGAGCAGGATGGAACCTTCACAAATG 253

DB 89 ATCAATCTACCAATTTTGTCCACAATAAGACAGAAAGATGGAACTCCCTTAAGGAT 148

QY 254 AAGCTGTGAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAAATAGATTACTGG 313

DB 149 CAAGGAGTAAATACCTCCACCTAAGCGGCTCTTTTACAGGACAAATAACCTATGG 208

QY 314 AAGCGGGAATCAAGCCAGTTTATGTTTGTGATGGCAAGCTCTCTGATATGAAGAAACAAG 373

DB 209 AGGCTGAATAAAACCTGTGTATGTTTGTGATGAGAACTCCAGAAATTCAAAAAGAAAG 268

QY 374 AGCTTGTAAAGTACTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGAGTAG 433

DB 269 AGCTGAAAAAAGAGAGAGGAGGAGGAGTGAAGAAAGTGGGAGAGAGCACTTG 328

QY 434 AGGTAGGAGATAAGATGCGATTGCAAAATTTGAGCAAGAGGACTGTAAAGTCAAGGC 493

DB 329 AAAAGAGAGATAGAGGAGCAAGAAATATGCCCAAGAGCAACAGGGTAAATGAA 388

QY 494 AACACAAGAGATTTGTAACGGCTATTAAAGACTTTATGGGGTTTCTGTTCTAGAGGAC 553

DB 389 TGCTCATCGAGGATGCAAAAAACTCTTAGAGCTTATGGGAATTCCTATAGTCAAGCAC 448

QY 554 CTCTGAGCAGAGCAGAAATGTGACGCCCTTTTCATATAAGATTAAGGTGTCGCTGTTG 613

DB 449 CTAGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 508

QY 614 CTTCAGAGATATGAGCTCCCTTACTTTTGGGGTTCACAGGTTTCTCTGCTCATTAATGG 673

DB 509 CTAGTCAAGATTAGGATTCCTTACTTTTGGAGTCCAGAGCTTTGTAGAACTTAACAA 568

QY 674 ATCCAGTTTCCAGAAATACCTGTGATGGAATTTGATGTTG 715

DB 569 TAACAGGAAAAAGAAAGTTGCTCGGGAAAAATGTCTACGTG 610

RESULT 11

US-08-823-516-78
; Sequence 78, Application US/08823516
; Patent No. 5994069

GENERAL INFORMATION:

; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.

;; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;; TITLE OF INVENTION: Sequential Invasive Cleavages
;; NUMBER OF SEQUENCES: 163
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States Of America
;; ZIP: 94104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/823,516
;; FILING DATE: 24-MAR-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US97/01072
;; FILING DATE: 21-JAN-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/759,038
;; FILING DATE: 02-DEC-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/758,314
;; FILING DATE: 02-DEC-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/756,386
;; FILING DATE: 29-NOV-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/682,853
;; FILING DATE: 12-JUL-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/599,491
;; FILING DATE: 24-JAN-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane E.
;; REGISTRATION NUMBER: 40,027
;; REFERENCE/DOCKET NUMBER: FORS-02736
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEQ ID NO: 78:
;; LENGTH: 1023 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1020
;; US-08-823-516-78

Query Match 10.9%; Score 159.6; DB 2; Length 1023;
Best Local Similarity 54.6%; Pred. No. 2e-32;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;
QY 134 CGATGAAGGAGCAGAGCTTCGAGGCTACTTCGGCCCGCAAAATCGCGCGCCAGCA 193
Db 29 CAGAAAAGAAATGGATTAGAAACCTATACGGAAAAATCGCAATCGAGCGCTTTA 88
QY 194 TGAGCATATACCAAGTTCCTGATTGTAGTTGGAAGGACAGGCTGGAACCTCTCACAATG 253
Db 89 ATGCAATCTACCAATTTTTCACACATAAGACAGAAAGATGGAACCTCCACTTATGATT 148
QY 254 AAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAAATAGATTAATG 313
Db 149 CAAAGGGTAGAATAAACCTCCACGCTAAGCGGGCTCTTTTACAGGACAAATAAACCTAATG 208
QY 314 AAGCGGAATCAAGCCAGTTTATGTTTTCATGGCAAGCCCTCTGATATGAAGAACAG 373

Db 209 AGGCTGGAATAAAACCTGTGTATGTTTGTGAGAACCTCCAGAAATTCAAAAGAAAG 268
QY 374 AGCTTGTCTAAAGACTACTCAAAAAGAGATGATCAACCAAGAGATCTGACTCAGGCAGTAG 433
Db 269 AGCTCGAAAAAAGAGAGAGACGAGAGAGGAGCTGAAGAAAAGTGGAGAGAACCACTTTG 328
QY 434 AGGTAGGAGATAAAGATCCGATTGAAAAAATTTGACGAAGAGGACTGTAAAGTCAACAGGC 493
Db 329 AAAAAGGAGAGATAGAGGAGCAAGCAAAATATGCCAAAGAGCAACCGGTAAATGAAA 388
QY 494 AACACAGGAAGATTGTAACCGGTATTAAAGCTATTAAAGCTATTGGGGGTTCCCTGTTTAGAGGCAC 553
Db 389 TGCTCATCGAGGATGCAAAAACCTCTTAGAGCTTATGGGAATTCCTATAGTTCGAAGCAC 448
QY 554 CTCTGGAAGCAGAGCAGAAATGTCAGCCCTTTTCATATAACGATAAGTGTTCCTGTTG 613
Db 449 CTACGAGGAGAGGCCCAAGCTGCATATATGGCCGCAAGAGGAGCGGTGATGCAATCGG 508
QY 614 CTTCAGAAGATATGGACTCCCTTACTTTTGGGGTCCACAGGTTCCCTTCGTCATTAAATGG 673
Db 509 CTAGTCAAGATTACGATTCCCTACTTTTGGAGCTCCAAGACTTTGTAGAAACTTAACAA 568
QY 674 ATCCAAGTTCACAGAAATACCTGTGATGGAATTTGATGTTG 715
Db 569 TAACAGGAAAAAGAAAGTTGCTCTGGGAAAAATGTTACGTGCG 610

RESULT 12
US-08-759-038-114
; Sequence 114, Application US/08759038
; Patent No. 6090543
; GENERAL INFORMATION:
; APPLICANT: Prudent, James R.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; APPLICANT: Dahlberg, James E.
; TITLE OF INVENTION: Cleavage Of Nucleic Acids
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,038
; FILING DATE: 02-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 114:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
US-08-759-038-114

Query Match
Best Local Similarity 10.9%; Score 159.6; DB 3; Length 1023;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAAGGAGCAGAGATTGCGAGAGCTACTCGGCCGCAAAATCGCGTCGACGCCAGCA 193
Db 29 CAAGAAAGAAATTTGATTAGAAAACCTATACGGGAAAAAAATCGCAATCGACCTCTTA 88

QY 194 TGAGCATATACCACTTCCTGATTGTAGTTCGAAGGACAGCATGGAACCTTCACAATG 253
Db 89 ATGCAATCTACCAATTTTGTCCACAATAAGACAGAAAGATGGAACCTTCACTTA7GGATT 148

QY 254 AAGCTGGTGAAGTCACTAGTCACTTTTCAAGGAATGTTCAACCGGACAAATAAGATTACTGG 313
Db 149 CAAGGGTAGAATAACCTCCACCTAAGCGGGCTCTTTACAGGACAAATAACCTTAATGG 208

QY 314 AAGCGGGAATCAAGCAGTTTATGTTTTGATGGCAAGCCTCTGTATATGAAGAAACAAG 373
Db 209 AGGCTGGAATAAAACCTGTGTATGTTTTGATGGAGAAAGCTGGAAGAAAGTGGAGAGAGCACTTG 268

QY 374 AGCTTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACTG...GCAGTAG 433
Db 269 AGCTCGAAAAAGAGAGAGGAGGAGAGAGCTGGAAGAAAGTGGAGAGAGCACTTG 328

QY 434 AGGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAAAGGTTCACAAGGC 493
Db 329 AAAAGGAGAGATAGAGGAGCAAGAAATATGCCCAAGAGCAACACAGGTAATGAA 388

QY 494 AACACAACGAAGATTGTAACCGGCTATTAGACTTTATGGGGTTTCTGTGTAGAGGCAC 553
Db 389 TGCTCATCGAGGATGCAAAAAAATCTTACAGCTTATGGGAATTCCTATATAGTTCAAGCAC 448

QY 554 CTTCTGAACGACAGCAGAAATGTGCAGCCCTTTGCATATAACGATAAGGTTCCTGCTTGG 613
Db 449 CTAGCGAGGAGAGCCCAAGCTGCATATATGCGCCGCAAGGGGAGCGTGTATGCATCGG 508

QY 614 CTTCAAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTCGTCAATTAATGG 673
Db 509 CTAGTCAAGATTACGATTCCCTACTTTTGGAGCTCCAAGACTTGTAGAAACTTAACAA 568

QY 674 ATCCAAGTTCCAAGAAATACCTGTGATGGAATTTGATGTTG 715
Db 569 TAACAGAAAAAGAAAGTTCCCTGGGAAAAATGTTCACTGCG 610

RESULT 13
US-08-758-314-114
; Sequence 114, Application US/08758314
; Patent No. 6090606
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Improved Cleavage Agents
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,314
; FILING DATE: 02-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02575
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
US-08-758-314-114

Query Match
Best Local Similarity 10.9%; Score 159.6; DB 3; Length 1023;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAAGGAGCAGAGATTGCGAGAGCTACTCGGCCGCAAAATCGCGTCGACGCCAGCA 193
Db 29 CAAGAAAGAAATTTGATTAGAAAACCTATACGGGAAAAAAATCGCAATCGACCTCTTA 88

QY 194 TGAGCATATACCACTTCCTGATTGTAGTTCGAAGGACAGCATGGAACCTTCACAATG 253
Db 89 ATGCAATCTACCAATTTTGTCCACAATAAGACAGAAAGATGGAACCTTCACTTA7GGATT 148

QY 254 AAGCTGGTGAAGTCACTAGTCACTTTTCAAGGAATGTTCAACCGGACAAATAAGATTACTGG 313
Db 149 CAAGGGTAGAATAACCTCCACCTAAGCGGGCTCTTTACAGGACAAATAACCTTAATGG 208

QY 314 AAGCGGGAATCAAGCAGTTTATGTTTTGATGGCAAGCCTCTGTATATGAAGAAACAAG 373
Db 209 AGGCTGGAATAAAACCTGTGTATGTTTTGATGGAGAAAGCTGGAAGAAAGTGGAGAGAGCACTTG 268

QY 374 AGCTTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACTG...GCAGTAG 433
Db 269 AGCTCGAAAAAGAGAGAGGAGGAGAGAGCTGGAAGAAAGTGGAGAGAGCACTTG 328

QY 434 AGGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAAAGGTTCACAAGGC 493
Db 329 AAAAGGAGAGATAGAGGAGCAAGAAATATGCCCAAGAGCAACACAGGTAATGAA 388

QY 494 AACACAACGAAGATTGTAACCGGCTATTAGACTTTATGGGGTTTCTGTGTAGAGGCAC 553
Db 389 TGCTCATCGAGGATGCAAAAAAATCTTACAGCTTATGGGAATTCCTATATAGTTCAAGCAC 448

QY 554 CTTCTGAACGACAGCAGAAATGTGCAGCCCTTTGCATATAACGATAAGGTTCCTGCTTGG 613
Db 449 CTAGCGAGGAGAGCCCAAGCTGCATATATGCGCCGCAAGGGGAGCGTGTATGCATCGG 508

QY 614 CTTCAAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTCGTCAATTAATGG 673
Db 509 CTAGTCAAGATTACGATTCCCTACTTTTGGAGCTCCAAGACTTGTAGAAACTTAACAA 568

QY 674 ATCCAAGTTCCAAGAAATACCTGTGATGGAATTTGATGTTG 715
Db 569 TAACAGAAAAAGAAAGTTCCCTGGGAAAAATGTTCACTGCG 610
```


GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 19:52:22 ; Search time 195 Seconds

(without alignments)
9323.473 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463

Sequence: 1 caccgagaatagctcgccgccc.....aaaaaaaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1463	100.0	1463	10	US-09-805-311-1
2	1421.8	97.2	1478	10	US-09-805-311-7
3	1409.8	96.4	1541	10	US-09-805-311-3
4	1350.2	92.3	1381	10	US-09-805-311-5
5	159.6	10.9	1023	9	US-10-033-297-78
6	159.6	10.9	1023	9	US-09-940-244-78
7	152.8	10.4	1054	9	US-09-940-244-281
8	151.4	10.3	485	9	US-09-918-995-30803
9	149.6	10.2	1115	9	US-09-940-244-289
10	139.6	9.5	510	10	US-09-864-761-16172
11	135.6	9.3	1032	9	US-09-940-244-336
12	133	9.1	1056	9	US-09-940-244-360
13	128	8.7	514	9	US-09-940-244-287
14	127.6	8.7	1023	9	US-09-940-244-383
15	126.4	8.6	1056	9	US-09-940-244-336
16	126.4	8.6	1164	9	US-09-940-244-272
17	125.6	8.6	1115	9	US-09-940-244-275
18	124	8.5	1041	9	US-09-940-244-378
19	121	8.3	1023	9	US-09-940-244-401

20	112.6	7.7	1056	9	US-09-940-244-340	Sequence 340, App
21	110.2	7.5	889	9	US-09-940-244-271	Sequence 271, App
22	108.2	7.4	1023	9	US-09-940-244-364	Sequence 364, App
23	108	7.4	1053	9	US-09-940-244-344	Sequence 344, App
24	106.2	7.3	981	9	US-09-940-244-352	Sequence 352, App
25	98.6	6.7	1011	9	US-09-940-244-164	Sequence 164, App
26	97.8	6.7	1008	9	US-09-940-244-348	Sequence 348, App
27	97.6	6.7	1164	9	US-09-940-244-286	Sequence 286, App
28	96	6.6	1062	9	US-09-940-244-374	Sequence 374, App
29	96	6.6	1071	9	US-09-940-244-388	Sequence 388, App
30	95.2	6.5	1017	9	US-09-940-244-369	Sequence 369, App
31	95.2	6.5	1017	10	US-09-777-430A-47	Sequence 47, Appl
32	94.8	6.5	386	9	US-09-940-244-276	Sequence 276, App
33	90	6.2	546	10	US-09-864-761-13411	Sequence 13411, A
34	89.2	6.1	981	9	US-10-033-297-74	Sequence 74, Appl
35	89.2	6.1	981	9	US-09-940-244-74	Sequence 32666, A
36	86.6	5.9	245	10	US-09-864-761-32666	Sequence 393, App
37	85.6	5.9	1053	9	US-09-940-244-393	Sequence 260, App
38	76.2	5.2	777	9	US-09-940-244-260	Sequence 274, App
39	74.6	5.1	840	9	US-09-940-244-274	Sequence 264, App
40	69.4	4.7	987	9	US-09-940-244-264	Sequence 273, App
41	67.8	4.6	296	9	US-09-940-244-273	Sequence 397, App
42	67.2	4.6	985	9	US-09-940-244-397	Sequence 282, App
43	66	4.5	514	9	US-09-940-244-282	Sequence 278, App
44	58	4.0	668	9	US-09-940-244-278	Sequence 4845, Ap
45	51.6	3.5	418	10	US-09-960-352-4845	

ALIGNMENTS

RESULT 1

US-09-805-311-1
; Sequence 1, Application US/09805311
; Patent No. US20010034886A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961D
; CURRENT APPLICATION NUMBER: US/09/805,311
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/426,557
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/112,332
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)...(1221)
US-09-805-311-1

Query Match	100.0%	Score 1463;	DB 10;	Length 1463;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1463;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CACGAGAAATAGCTCGCGCGCGGGTTTCTTGGGCACATCCGGCTCAGCGCGCGCGGCCA	60	
Db	1	CACGAGAAATAGCTCGCGCGCGGGTTTCTTGGGCACATCCGGCTCAGCGCGCGCGGCCA	60	
Qy	61	CCCGCCACAGCGCGCGAGAGATGGGATCAAGGGTTTGACGAAACTCTCTGGCGGAC	120	
Db	61	CCCGCCACAGCGCGCGAGAGATGGGATCAAGGGTTTGACGAAACTCTCTGGCGGAC	120	
....Qy	121	AATCGGCCAAGCGATGAAGSACAGAGTTCAGAGCTACTTCGCGCGCAATCGCC	180	
Db	121	AATCGGCCAAGCGATGAAGSACAGAGTTCAGAGCTACTTCGCGCGCAATCGCC	180	

```
QY 181 GTCCAGCCGACGATGAGCATATACACAGTTCTCTGATTTGGAAGCAGAGCATGAA 240
DB 181 GTCCAGCCGACGATGAGCATATACACAGTTCTCTGATTTGGAAGCAGAGCATGAA 240
QY 241 ACTCTCAAAATGAAGCTGGTGAAGTCTAGTCAATTTGCAAGGAATGTTCAACCGGACA 300
DB 241 ACTCTCAAAATGAAGCTGGTGAAGTCTAGTCAATTTGCAAGGAATGTTCAACCGGACA 300
QY 301 ATAAGATTACTGGAAGCGGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCCTCCTGAT 360
DB 301 ATAAGATTACTGGAAGCGGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCCTCCTGAT 360
QY 361 ATGAAGAAACAGAGCTTGCTTAAAGATACTCAAAAGAGATGATGCAACCAAGATCTG 420
DB 361 ATGAAGAAACAGAGCTTGCTTAAAGATACTCAAAAGAGATGATGCAACCAAGATCTG 420
QY 421 ACTGAGGCAGTAGAGGTAGGAGATAAGATGCGATTGAAAAATTGAGCAAGAGGACTGTA 480
DB 421 ACTGAGGCAGTAGAGGTAGGAGATAAGATGCGATTGAAAAATTGAGCAAGAGGACTGTA 480
QY 481 AAGTTCACAAAGCACAACAGAGATTGTAACGGCTATTAAAGACTATGGGGGTTCCCT 540
DB 481 AAGTTCACAAAGCACAACAGAGATTGTAACGGCTATTAAAGACTATGGGGGTTCCCT 540
QY 541 GTTGTAGAGGCACCTTCTGAAGCAGAGCAGAAATGTGAGCCCTTTGCATAAAGCATAAG 600
DB 541 GTTGTAGAGGCACCTTCTGAAGCAGAGCAGAAATGTGAGCCCTTTGCATAAAGCATAAG 600
QY 601 GTGTTGCGTGTGCTTCAGAGATATGACCTCCCTTACTTTTGGGGCTCCACGGTTCCCT 660
DB 601 GTGTTGCGTGTGCTTCAGAGATATGACCTCCCTTACTTTTGGGGCTCCACGGTTCCCT 660
QY 661 CGTCTATTAAATGGATCCAAAGTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAG 720
DB 661 CGTCTATTAAATGGATCCAAAGTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAG 720
QY 721 GTTGTGAGGAGCTGAACTCACCATGACAGCTTCATTTGATGTCATCCTGTGGA 780
DB 721 GTTGTGAGGAGCTGAACTCACCATGACAGCTTCATTTGATGTCATCCTGTGGA 780
QY 781 TGTGACTATTGTGATGATCAATCAAGGATATCGGGGGCAACAGCTCTGAACTTTATTCGT 840
DB 781 TGTGACTATTGTGATGATCAATCAAGGATATCGGGGGCAACAGCTCTGAACTTTATTCGT 840
QY 841 CAACATGGTCCATAGAAGCATCTTGAGATCTTATTAAGACAGATATCAAAATTCCT 900
DB 841 CAACATGGTCCATAGAAGCATCTTGAGATCTTATTAAGACAGATATCAAAATTCCT 900
QY 901 GAGGACTGGCCTTACCAAGAGCTCGAGCCTTGTTCAAGGAGCCTTAATGTCACATTTGGAT 960
DB 901 GAGGACTGGCCTTACCAAGAGCTCGAGCCTTGTTCAAGGAGCCTTAATGTCACATTTGGAT 960
QY 961 ATTCCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCAATGTTTCCCTGGTAAA 1020
DB 961 ATTCCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCAATGTTTCCCTGGTAAA 1020
QY 1021 GATAATGTTTCAAGAAAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAG 1080
DB 1021 GATAATGTTTCAAGAAAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAG 1080
QY 1081 AATAATGCTGCAAGGAGACTCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCA 1140
DB 1081 AATAATGCTGCAAGGAGACTCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCA 1140
QY 1141 CGCGTAAACCGGAGGACTTCGATATAAAACAAGCAGGAGCTGGCAACGAAGAAACA 1200
DB 1141 CGCGTAAACCGGAGGACTTCGATATAAAACAAGCAGGAGCTGGCAACGAAGAAACA 1200
QY 1201 AAGGCTGGTGAAGAGAAATATCTTGGATGCTTGTATGACACTACGACTACGAAAG 1260
DB 1201 AAGGCTGGTGAAGAGAAATATCTTGGATGCTTGTATGACACTACGACTACGAAAG 1260
QY 1261 CAGCGGTGGCGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTT 1320
```

```
DB 1261 CAGCGGTGGCGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTT 1320
QY 1321 GGTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAAGTTAGTTAGTTGTTGAAGAGATTGGTG 1380
DB 1321 GGTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAAGTTAGTTAGTTGTTGAAGAGATTGGTG 1380
QY 1381 TACCAAGTAAACAAACTTATCGCTGTTTTTTTACTTCTTCTCTTGAAGTAAAAA 1440
DB 1381 TACCAAGTAAACAAACTTATCGCTGTTTTTTTACTTCTTCTCTTGAAGTAAAAA 1440
QY 1441 AAAAAA 1463
DB 1441 AAAAAA 1463
```

RESULT 2

```
US-09-805-311-7
; Sequence 7, Application US/09805311
; Patent No. US20010034886A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961D
; CURRENT APPLICATION NUMBER: US/09/805,311
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/426,557
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/112,332
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1478
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)...(1233)
US-09-805-311-7
```

Query Match 97.2%; Score 1421.8; DB 10; Length 1478;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```
QY 7 ATATGCTCGGCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCGCCGCCACCCGCC 66
DB 19 ATATGCTCGGCTCGCGGTTCTTGGCCACTCCGGCTCAGCGCGCGCCGCCACCCGCC 78
QY 67 ACAGCGCGCGCAGACGAGATGGCATCAAGGTTTGGAGAACTGCTGGCGACAAATCGG 126
DB 79 ACAGCGCGCGCAGACGAGATGGCATCAAGGTTTGGAGAACTGCTGGCGACAAATCGG 138
QY 127 CCCAAGCGATGAAGGAGCAGAAAGTTTCGAGACTACTTGGCGCGCAAAATGCCGTCGAC 186
DB 139 CCCAAGCGCATGAAGGAGCAGAAAGTTTCGAGACTACTTGGCGCGCAAAATGCCGTCGAC 198
QY 187 GCCAGCATGAGCATATACCAGTTCTCTGATGTTGGAAGCAGAGCATGGAATCTC 246
DB 199 GCCAGCATGAGCATATACCAGTTCTCTGATGTTGGAAGCAGAGCATGGAATCTC 258
QY 247 ACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACGGCAATAAGA 306
DB 259 ACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACGGCAATAAGA 318
QY 307 TTAATGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGGCAAGCTCTCTGATATGAG 366
DB 319 TTAATGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGGCAAGCTCTCTGATATGAG 378
QY 367 AAACAAGAGCTTGTAAAGATACACTCAAAAGAGATGATGCAACCAAGATCTGACTGAG 426
DB 379 AAACAAGAGCTTGTAAAGATACACTCAAAAGAGATGATGCAACCAAGATCTGACTGAG 438
```

QY	427	GCAGTAGAGGTAGAGATAAAGATGCGGATGAAATTTGACCAAGAGACTGTAAAGTC	486
DB	439	GCAGTAGAGGTAGAGATAAAGATGCGGATGAAATTTGACCAAGAGACTGTAAAGTC	498
QY	487	ACAAGGCAACACACAGAGATGTAAGCGGTATTAAGACTTATGGGGTTCCTGTTGTA	546
DB	499	ACAAGGCAACACACAGAGATGTAAGCGGTATTAAGACTTATGGGGTTCCTGTTGTA	558
QY	547	GAGGCACCTTCTGAAGCAGACAGAGATGTCAGCCCTTTGCATAAAGCATAAAGTGTC	606
DB	559	GAGGCACCTTCTGAAGCAGACAGAGATGTCAGCCCTTTGCATAAAGCATAAAGTGTC	618
QY	607	GCTGTGCTCAGAAAGATATGAGTCCCTTACTTTTGGGGTCCACGGTTCCTTCTGTCAT	666
DB	619	GCTGTGCTCAGAAAGATATGAGTCCCTTACTTTTGGGGTCCACGGTTCCTTCTGTCAT	678
QY	667	TTAATGGATCAAGTTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTG	726
DB	679	TTAATGGATCAAGTTCCAAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTG	738
QY	727	GAGGAGCTTGAACCTCACCATGACAGTTCATTTGATTTGTCATCCTGTGTGGATGTGAC	786
DB	739	GAGGAGCTTGAACCTCACCATGACAGTTCATTTGATTTGTCATCCTGTGTGGATGTGAC	798
QY	787	TATTGTGATAGCATCAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAACAT	846
DB	799	TATTGTGATAGCATCAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAACAT	858
QY	847	GGGTCCATAGAAGCATCTTGAGAACTTTAATAAGACAGATATCAAAATTCCTGAGGAC	906
DB	859	GGGTCCATAGAAGCATCTTGAGAACTTTAATAAGACAGATATCAAAATTCCTGAGGAC	918
QY	907	TGGCCCTTACCAAGAAGCTCGACGCTTGTTCAAGGAGCCTAATGTGCATTTGATTTCT	966
DB	919	TGGCCCTTACCAAGAAGCTCGACGCTTGTTCAAGGAGCCTAATGTGCATTTGATTTCT	978
QY	967	GAGCTAAATGAGCTGCACCTCATGAGGAGGCTCTCATAAAGTTTCTGCTTAAAGATAAT	1026
DB	979	GAGCTAAATGAGCTGCACCTCATGAGGAGGCTCTCATAAAGTTTCTGCTTAAAGATAAT	1038
QY	1027	GGTTTCAAGCAGAGATCGGTGCAAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAATAA	1086
DB	1039	GGTTTCAATGAAGATCGGTGCAAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAATAA	1098
QY	1087	TGTCGCAAGGAGACTCGAGTCTTTTCAAGCCCACTGCCACCATCAGCACCGCTA	1146
DB	1099	TGTCGCAAGGAGACTCGAGTCTTTTCAAGCCCACTGCCACCATCAGCACCGCTA	1158
QY	1147	AAACGGAAGGAGACTTCGGATAAAACAAGCAGCCAGCTGCGAACAAGAAACAAAGGCT	1206
DB	1159	AAACGGAAGGAGACTTCGGATAAAACAAGCAGCCAGCTGCGAACAAGAAACAAAGGCT	1218
QY	1207	GGTGGAAAGAAATAATCTTGGATGCTTGATGTACAACTAGCACTACCAAGCAGCGG	1266
DB	1219	GGTGGAAAGAAATAATCTTGGATGCTTGATGTACAACTAGCACTACCAAGCAGCGG	1278
QY	1267	TGCGGTGATCAGTTCGCTTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTTGTTAAA	1326
DB	1279	TGCGGTGATCAGTTCGCTTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTTGTTAAA	1338
QY	1327	AGTTTGTCTCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGATTTGGTGTACCAA	1386
DB	1339	AGTTTGTCTCATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGATTTGGTGTACCAA	1398
QY	1387	GTAACAAACTTATCGCTGTTTTTACTTCTTGTCCTTTGAGTAAACCAAGGAGGAGGAGG	1446
DB	1399	GTAACAAACTTATCGCTGTTTTTACTTCTTGTCCTTTGAGTAAACCAAGGAGGAGGAGG	1458
QY	1447	AAAAAAAAAAAAAAAAAAAA 1463	
DB	1459	AAAAAAAAAAAAAAAAAAAA 1475	

RESULT 3									
US-09-805-311-3									
; Sequence 3, Application US/09805311									
; Patent No. US20010034886A1									
; GENERAL INFORMATION:									
; APPLICANT: Mahajan, Pramod B.									
; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses									
; TITLE OF INVENTION: Thereof									
; FILE REFERENCE: 0961D									
; CURRENT APPLICATION NUMBER: US/09/805,311									
; CURRENT FILING DATE: 2001-03-13									
; PRIOR APPLICATION NUMBER: 09/426,557									
; PRIOR FILING DATE: 1999-10-22									
; PRIOR APPLICATION NUMBER: 60/112,332									
; PRIOR FILING DATE: 1998-12-15									
; NUMBER OF SEQ ID NOS: 10									
; SOFTWARE: FastSeq for Windows Version 3.0									
; SEQ ID NO 3									
; LENGTH: 1541									
; TYPE: DNA									
; ORGANISM: Zea mays									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (79)...(1215)									
US-09-805-311-3									
Query Match 96.4%; Score 1409.8; DB 10; Length 1541;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 1411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	19	CGCGGGTTCTTCGGCCACTCCGGCTCAGCGCGCGCGCGCGCGCGCGCGCA	78						
DB	13	CGCGGGTTCTTCGGCCACTCCGGCTCAGCGCGCGCGCGCGCGCGCGCGCA	72						
QY	79	GACGAGATGGGATCAAGGGTTTGACGAACTGCTGCGGACAAATGCGCCAAAGCGATG	138						
DB	73	GACGAGATGGGATCAAGGGTTTGACGAACTGCTGCGGACAAATGCGCCAAAGCGATG	132						
QY	139	AAGGACGAGAAGTTCGAGAGCTACTTCGCGCGCAAAATCGCGTCGACGCCAGCATGAGC	198						
DB	133	AAGGACGAGAAGTTCGAGAGCTACTTCGCGCGCAAAATCGCGTCGACGCCAGCATGAGC	192						
QY	199	ATATACCAAGTTCCTGATTTGATTTGGAAGGACAGGCGATGGAACCTCTCACAATGAAGCT	258						
DB	193	ATATACCAAGTTCCTGATTTGATTTGGAAGGACAGGCGATGGAACCTCTCACAATGAAGCT	252						
QY	259	GGTGAAGTCACTAGTTCATTTGCAAGGAATGTTCAACCGGACAAATAAGATTACTTGAAGCG	318						
DB	253	GGTGAAGTCACTAGTTCATTTGCAAGGAATGTTCAACCGGACAAATAAGATTACTTGAAGCG	312						
QY	319	GGATCAAGCCAGTTTATGTTTGTGATGGCAAGCTCCTGATATGAAGAAACAAGAGCTT	378						
DB	313	GGATCAAGCCAGTTTATGTTTGTGATGGCAAGCTCCTGATATGAAGAAACAAGAGCTT	372						
QY	379	GCTAAAGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGCGAGTAGAGTA	438						
DB	373	GCTAAAGATACTCAAAAAGAGATGATCAACCAAGATCTGACTGAGGCGAGTAGAGTA	432						
QY	439	GGAGATAAAGATGCGATTGAAAAATTTGACGAGAGGACTGTAAAGGTCACAAAGCAACAC	498						
DB	433	GGAGATAAAGATGCGATTGAAAAATTTGACGAGAGGACTGTAAAGGTCACAAAGCAACAC	492						
QY	499	AACGAAGATTGTAACCGCTATTAAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCT	558						
DB	493	AACGAAGATTGTAACCGCTATTAAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCT	552						
QY	559	GAAGCAGAAGCAGAAATGTCAGCCCTTTGCATAAACGATAAGGTTTTCGCTGTTCCTCA	618						
DB	553	GAAGCAGAAGCAGAAATGTCAGCCCTTTGCATAAACGATAAGGTTTTCGCTGTTCCTCA	612						
QY	619	GAAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCATTTTATGGATCA	678						

Db 901 GTACATTGATATTCCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCATAGT 960
Qy 1009 TTCCTGGTAAAGATAATGGTTTCAACGAGATCGGGTGACAAAGCCATAGAGATC 1068
Db 961 TTCCTGGTAAAGATAATGGTTTCAACGAGATCGGGTGACAAAGCCATAGAGATC 1020
Qy 1069 AAATCTGCCAAGAATAAATCGTCGCAAGGAAGACTCGAGTCCTTTTCAAGCCAACTGCC 1128
Db 1021 AAATCTGCCAAGAATAAATCGTCGCAAGGAAGACTCGAGTCCTTTTCAAGCCAACTGCC 1080
Qy 1129 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGATATAAACAAGCAGGAGCTGCC 1188
Db 1081 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGATATAAACAAGCAGGAGCTGCC 1140
Qy 1189 ACAAGAATAAAGAGCTGGTGGAAAGAGAAATAATCTTGGATGCTTGATGATACACTA 1248
Db 1141 ACAAGAATAAAGAGCTGGTGGAAAGAGAAATAATCTTGGATGCTTGATGATACACTA 1200
Qy 1249 CGACTACGAAAGCAGCGGTGGCTGATCACTTCGCTTAGATATTTAACTCCCTGTTTTA 1308
Db 1201 CGACTACGAAAGCAGCGGTGGCTGATCACTTCGCTTAGATATTTAACTCCCTGTTTTA 1260
Qy 1309 ACTCAGAGCTTTGGTAAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTGTTG 1368
Db 1261 ACTCAGAGCTTTGGTAAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTGTTG 1320
Qy 1369 AGAGATTGGTGACCAAGTAACAACACTTATCGCTGTTTTT 1411
Db 1321 AGAGATTGGTGACCAAGTAACAACACTTATCGCTGTTTTT 1363

RESULT 5

US-10-033-297-78
; Sequence 78, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:

APPLICANT: Hall, Jeff G.
Mast, Andrea L.
Brow, Mary Ann D.

TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages

NUMBER OF SEQUENCES: 163

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
Zip: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/033.297

FILING DATE: 12-NOV-1999

CLASSIFICATION DATA:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/350.597

FILING DATE: 09-Jul-1999

APPLICATION NUMBER: US/08/823.516

FILING DATE: 24-MAR-1997

APPLICATION NUMBER: PCT/JUS97/01072

FILING DATE: 21-JAN-1997

APPLICATION NUMBER: US 08/759,038

FILING DATE: 02-DEC-1996

APPLICATION NUMBER: US 08/758,314

FILING DATE: 02-DEC-1996

APPLICATION NUMBER: US 08/756,386

FILING DATE: 29-NOV-1996

APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1020
SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-10-033-297-78

Query Match 10.9%; Score 159.6; DB 9; Length 1023;
Best Local Similarity 54.6%; Pred. No. 2.4e-34;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

Qy 134 CGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCGTCGAGCCAGCA 193
Db 29 CAAGAAAAAGAAATTCAGTTAGAAAACTATACGGGAAAAAAATCGCAATCGACGCTT 88
Qy 194 TGAGCATATACCAAGTTCCTGATTGTAGTTGGAAGGACAGGCGATGGAACCTCTCACA 253
Db 89 ATGCAATCTACCAATTTTGTCCACAAATAGACAGAAAGATGCAACTCCACTTATGG 148
Qy 254 AAGCTGGTGAAGTCACTAGTCACTTTTCAGAGGAATGTTCAACCGGAGCAATTAAG 313
Db 149 CAAAGGGTAGAATAACCTCCACCTTAAGCGGGCTCTTTTACAGGACAAATTAACCT 208
Qy 314 AAGCGGGATCAAGCAGTTTATGTTTTCATGCGCAAGCTCTCTGATATCAAGAAACA 373
Db 209 AGGCTGGAATAAAACCTGTGTATGTTTTCATGCGCAAGCTCTCTGATATCAAGAA 268
Qy 374 AGCTTGTCTAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTCAGCAGT 433
Db 269 AGCTCGAAAAAGAGAGAGCGGAGAGAGAACTGAAGAAAAAGTGGAGAGAACCTTG 328
Qy 434 AGGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAAAGTCAACA 493
Db 329 AAAAAAGGAGAGATAGAGGAAGCAAGAAAAATATGCCCAAGAGCAACCCAGGTA 388
Qy 494 AACACAACCAAGATTCTAAACGGCTATTAGACTTTATGGGGTTTCCTGTTGTAGAG 553
Db 389 TGCTCATCGAGGATGCAAAAAAATCTTTAGAGCTTTATGGGAATTCCTATAGTTC 448
Qy 554 CTCTTGAAGCAGAAGCAGAAATGTGCAGCCCTTTTGCATATAAAGCAAGATGCTCT 613
Db 449 CTAGCGAGGAGAGGCGCCCAAGCTGCATATATGCCCGCAAGGGGAGCGTGTATC 508
Qy 614 CTTCAAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCATTATGG 673
Db 509 CTAGTCAAGATTACGATTCCCTACTTTTGGAGCTCCCAAGACTTGTGTAGAACTTA 568
Qy 674 ATCCCAAGTTCCCAAGAAATACCTGTGATGCAATTTGATGTTG 715
Db 569 TAACAGGAAAAAGAAAGTTGCCTTGGGAAAAAATGTCTACGTCG 610

RESULT 6

US-09-940-244-78
; Sequence 78, Application US/09940244
; Publication No. US20030044796A1

; GENERAL INFORMATION:

; APPLICANT: Neri, Bruce P.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor
; APPLICANT: Smith, Lloyd M.
; TITLE OF INVENTION: Reactions on Dendrimers
; FILE REFERENCE: FORS-06478
; CURRENT APPLICATION NUMBER: US/09/940,244
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1020)
; OTHER INFORMATION:
US-09-940-244-78

Query Match 10.9%; Score 159.6; DB 9; Length 1023;
Best Local Similarity 54.6%; Pred. No. 2.4e-34;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;
QY 134 CGATGAGGAGCAGAGTTCGAGAGCTACTTCGGCCGCCAAATCGCGTCGAGCCGACGA 193
Db 29 CAAGAAAGAAATTTGAGTTAGAAAACCTATACGGGAAAAATTCGCAATCGAGCTCTTA 88
QY 194 TGAGCATATACCACTTCTCTGATTTGAGGAGCAGCATGGAACCTCTCACAATG 253
Db 89 ATGCATCTACCAATTTTGTCCACATAAGACAGAAAGATGGAATCCACTTATGGATT 148
QY 254 AAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATTTCAACCGGACAAATAGATTACTGG 313
Db 149 CAAAGGTAGATAAATCCCTCCACCTTAAGCGGCTCTTTTACAGGACAAATAACTTATGG 208
QY 314 AAGCGGAATCAAGCAGTTTATGTTTGTGATGCAAGCTCTCTGATATGAAGAAACAAG 373
Db 209 AGCTGGAATFAAACCTGTGTATGTTTGTGATGGAACCTCCAGAAATTCAAAAAGAAAG 268
QY 374 AGCTTGCTAAAGATCTCAAAAGAGATGATCAACCAAGATCTGACTGAGGCGACTAG 433
Db 269 AGCTCGAAAAAGAGAGAGAGGAGAGAGAGCTGAAGAAAAGTGGAGAGAGCACTTG 328
QY 434 AGTAGAGATAAAGATGCGATTGAAAAATTTGACAGAGAGGACTGTAAAGGTCAACAAGC 493
Db 329 AAAAAAGAGAGATAGAGGAAGCAAGAAAAATATGCCCAAGAGCAACACGAGGTAAATGAA 388
QY 494 AACACAACGAGAGATTGTAACGCTATTAAAGCTTATGAGGAGGCTTATGGGGTTCCTGTTGAGAGGCAC 553
Db 389 TGCTCATCGAGGATGCAAAAAAACTCTTAGAGCTTATGGGAATTCCTATAGTTTCAAGCAC 448
QY 554 CTCTCTGAAGCAGAGCAGAAATGTCAGCCCTTTGCAATAAAGATAAGGTGTCGCTGTTG 613
Db 449 CTAGCGAGGAGAGGCGCCAGCTGATATAGCCCAAGAGGAGGCGGTGATGATCGG 508
QY 614 CTTCAAGAGATATGAGTCCCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 673
Db 509 CTAGTCAAGATTACGATCCCTACTTTTGGAGCTCCAAGCTTGTAGAACTTAAACAA 568
QY 674 ATCCAAGTTCCAGAAATACCTGTGATGAATTTGATGTTG 715
Db 569 TACAGGAAAAAGAAAGTTGCTGGGAAAAATGCTACGTCG 610

RESULT 7

US-09-940-244-281
; Sequence 281, Application US/09940244
; Publication No. US20030044796A1
; GENERAL INFORMATION:
; APPLICANT: Neri, Bruce P.
; APPLICANT: Hall, Jeff G.

; APPLICANT: Lyamichev, Victor
; APPLICANT: Smith, Lloyd M.
; TITLE OF INVENTION: Reactions on Dendrimers
; FILE REFERENCE: FORS-06478
; CURRENT APPLICATION NUMBER: US/09/940,244
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 281
; LENGTH: 1054
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-940-244-281

Query Match 10.4%; Score 152.8; DB 9; Length 1054;
Best Local Similarity 53.4%; Pred. No. 2e-32;
Matches 350; Conservative 0; Mismatches 297; Indels 9; Gaps 1;
QY 134 CGATGAGGAGCAGAGTTCGAGAGCTACTTCGGCCGCCAAATCGCGTCGAGCCGACGA 193
Db 73 CAAGAAAGAAATTTGAGTTAGAAAACCTATACGGGAAAAATTCGCAATCGAGCTCTTA 132
QY 194 TGAGCATATACCACTTCTCTGATTTGAGGAGCAGCATGGAACCTCTCACAATG 253
Db 133 ATGCAATCTACCAATTTTGTCCACATAAGACAGAAAGATGGAATCCACTTATGGATT 192
QY 254 AAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATTTCAACCGGACAAATAGATTACTGG 313
Db 193 CAAAGGTAGATAAATCCCTCCACCTTAAGCGGCTCTTTTACAGGACAAATAACTTATGG 252
QY 314 AAGCGGAATCAAGCAGTTTATGTTTGTGCAAGCTCTCTGATATGAAGAAACAAG 373
Db 253 AGCTGGAATFAAACCTGTGTATGTTTGTGATGGAACCTCCAGATTCGAAAAAGAAAG 312
QY 374 AGTTGCTAAAGATCTCAAAAGAGATGATCAACCAAGATCTGACTGAGGCGACTAG 433
Db 313 AGCTGAAAAAGAGAGAGAGAGAGAGAGCTGAAGAAAAAGTGGAGAGAGCACTTG 372
QY 434 AGCTAGGAGATAAGATGCGATTGAAAAATTTGACAGAGAGGACTGTAAAGTCAACAAGC 493
Db 373 AAAAAAGAGAGATAGAGAAAGCTGAAGTATGCAAGAGGGTTAGCTATCACTCCGA 432
QY 494 AACACAACGAGAGATTGTAACGCTATTAAAGCTTATGGGGTTCCTGTTGAGAGGCAC 553
Db 433 AATGTTGAAAACTCAATATTTGTTAGTTTGTAGGCGATTCCTGATGTTGAAGCTC 492
QY 554 CTCTCTGAAGCAGAGCAGAAATGTCAGCCCTTTGCAATAAAGATAAGGTGTCGCTGTTG 613
Db 493 CCTCTGAGGAGAGGCACAAAGCAAGCTATATGCAAAAGAGGAGATGTTTGGGCGATTG 552
QY 614 CTTCAAGAGATATGAGTCCCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 673
Db 553 TAAGTCAAGATTATGATGCTTGTATATGAGGCTCCGAGAGTTGTAGAAATTTAAACA 612
QY 674 ATCCAAGTTCCAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTTCGAGGAGC 733
Db 613 CTACAAAGAGAGATGCCAGAACTTA-----TTGAATTAATGAGTTTATAGAGATT 663
QY 734 TTGAAGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789
Db 664 TAAGAAATTTCTTGGATGATTTGATAGATATAGCCATATTTATGGGAAGTACTAT 719

RESULT 8

US-09-918-995-30803
; Sequence 30803, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES

QY	592	AACGATAAGGTGTTCCGCTGTTGCTTCAGAAAGATATGGACTCCCTTACTCTTTGGGGCTCCA	651
Db	493	CAAGGGTATACTTTGGCAGCGGCTACGCCAAGATACGATCTCTGCTTTTTGGTGCAAT	552
QY	652	CGGTTCCCTCGTCATTTAATGATCCAGTCCAGAAAATACCTGTGATGGAATTGAT	711
Db	553	AAATAATTAGAACTTAACATTAACTGGAAGAGGAAATTACCTAAAAAGACGTATAT	612
QY	712	GTTG	715
Db	613	GTAG	616

Search completed: May 8, 2003, 21:39:33
Job time : 200 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	1425.8	97.5	1552	11	AY104316 Zea mays
2	778.8	53.2	901	12	BG837708 Zm10_01r10
3	614.8	42.0	638	14	B0037559
4	560.8	38.3	586	9	AI181599
5	536.4	36.7	550	10	BE639422
6	534.4	36.5	554	10	BE639421
7	534.4	36.5	554	10	BE639422
8	534.4	36.5	554	10	BE639421
9	534.4	36.5	554	10	BE639422
10	534.4	36.5	554	10	BE639421
11	534.4	36.5	554	10	BE639422
12	534.4	36.5	554	10	BE639421
13	534.4	36.5	554	10	BE639422
14	534.4	36.5	554	10	BE639421
15	534.4	36.5	554	10	BE639422
16	534.4	36.5	554	10	BE639421
17	534.4	36.5	554	10	BE639422
18	534.4	36.5	554	10	BE639421
19	534.4	36.5	554	10	BE639422
20	534.4	36.5	554	10	BE639421
21	534.4	36.5	554	10	BE639422
22	534.4	36.5	554	10	BE639421
23	534.4	36.5	554	10	BE639422
24	534.4	36.5	554	10	BE639421
25	534.4	36.5	554	10	BE639422
26	534.4	36.5	554	10	BE639421
27	534.4	36.5	554	10	BE639422
28	534.4	36.5	554	10	BE639421
29	534.4	36.5	554	10	BE639422
30	534.4	36.5	554	10	BE639421
31	534.4	36.5	554	10	BE639422
32	534.4	36.5	554	10	BE639421
33	534.4	36.5	554	10	BE639422
34	534.4	36.5	554	10	BE639421
35	534.4	36.5	554	10	BE639422
36	534.4	36.5	554	10	BE639421
37	534.4	36.5	554	10	BE639422
38	534.4	36.5	554	10	BE639421
39	534.4	36.5	554	10	BE639422
40	534.4	36.5	554	10	BE639421
41	534.4	36.5	554	10	BE639422
42	534.4	36.5	554	10	BE639421
43	534.4	36.5	554	10	BE639422
44	534.4	36.5	554	10	BE639421
45	534.4	36.5	554	10	BE639422
46	534.4	36.5	554	10	BE639421
47	534.4	36.5	554	10	BE639422
48	534.4	36.5	554	10	BE639421
49	534.4	36.5	554	10	BE639422
50	534.4	36.5	554	10	BE639421
51	534.4	36.5	554	10	BE639422
52	534.4	36.5	554	10	BE639421
53	534.4	36.5	554	10	BE639422
54	534.4	36.5	554	10	BE639421
55	534.4	36.5	554	10	BE639422
56	534.4	36.5	554	10	BE639421
57	534.4	36.5	554	10	BE639422
58	534.4	36.5	554	10	BE639421
59	534.4	36.5	554	10	BE639422
60	534.4	36.5	554	10	BE639421
61	534.4	36.5	554	10	BE639422

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

/clone.lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
477 a 310 c 381 g 384 t
BASE COUNT
ORIGIN

```

Query Match	97.5%	Score	1425.8	DB	11	Length	1552
Best Local Similarity	99.9%	Prod. No.	8.7e-252				
Matches 1427	Conservative	0	Mismatches	2	Indels	0	Gaps
QY	3	CGAGAAATAGCTCGCGCGCGCGTGTCTTTGCGCCACTCGCGCTCGACCGCGCGCGCCAC	62				
DB	8	CGGAATATAGCTCGCGCGCGCGGTTCCTTTGCGCCACTCGCGCTCGACCGCGCGCGCCAC	67				
QY	63	CGCCACAGCCCGCGAGAGATGGGCATCAAGGGTTTGACGAATCTGCTGGCGGACAA	122				
DB	68	CGCCACAGCCCGCGAGAGATGGGCATCAAGGGTTTGACGAATCTGCTGGCGGACAA	127				
QY	123	TCCGCCCAAGCGGATGAAGAGCAGAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGGT	182				
DB	138	TGCGCCCAAGCGGATGAAGAGCAGAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGGT	187				
QY	183	CGAGCCAGCATGAGCATATACCACTTCTCATTTAGTTGGAAGGACAGCATGGAAC	242				
DB	188	CGAGCCAGCATGAGCATATACCACTTCTCATTTAGTTGGAAGGACAGCATGGAAC	247				
QY	243	TCTCACAATGAAGCTGGTGAAGTCACCTAGTCATTTGCAAGGAATGTTCAACCGGACAAT	302				
DB	248	TCTCACAATGAAGCTGGTGAAGTCACCTAGTCATTTGCAAGGAATGTTCAACCGGACAAT	307				
QY	303	AAGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCTCTCTGATAT	362				
DB	308	AAGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCTCTCTGATAT	367				
QY	363	GAAGAAACAAGAGCTTGCTAAAGATACTCAAAAGAGATGATCAACCAAGATCTGAC	422				
DB	368	GAAGAAACAAGAGCTTGCTAAAGATACTCAAAAGAGATGATCAACCAAGATCTGAC	427				
QY	423	TGAGGCAGTAGAGTAGAGATAAAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAA	482				
DB	428	TGAGGCAGTAGAGTAGAGATAAAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAA	487				
QY	483	GGTCAACAGGCAACACACGAAGATCTTAACCGCTATTAAAGACTTATGGGGTTCCTGT	542				
DB	488	GGTCAACAGGCAACACACGAAGATCTTAACCGCTATTAAAGACTTATGGGGTTCCTGT	547				
QY	543	TGTAGAGCCACCTTCTGAAGCAGAAGCAGAATGTGACGCCCTTTGCAATAACGATAGGT	602				
DB	548	TGTAGAGCCACCTTCTGAAGCAGAAGCAGAATGTGACGCCCTTTGCAATAACGATAGGT	607				
QY	603	GTTGCGCTGTTGCTTCAGAAAGATATGGACTCCCTACTTTTGGGGCTCCACGGTTCCTTCG	662				
DB	608	GTTGCGCTGTTGCTTCAGAAAGATATGGACTCCCTACTTTTGGGGCTCCACGGTTCCTTCG	667				
QY	663	TCATTTAATGGATCCAAGTTCCTCAAGAAAAATACCTGTGATGGAATTTGATCTCCCAAGGT	722				
DB	668	TCATTTAATGGATCCAAGTTCCTCAAGAAAAATACCTGTGATGGAATTTGATCTCCCAAGGT	727				
QY	723	TTTGGGAGAGCTTGAACCTACCATGGAAGAGTTCAATTTGATGATCCTGTGTGGATG	782				
DB	728	TTTGGGAGAGCTTGAACCTACCATGGAAGAGTTCAATTTGATGATCCTGTGTGGATG	787				
QY	783	TGACTATTGTGATAGCATCAAGGTTATCGGGGGCAACAGCTCTGAACCTTATTCGTCA	842				
DB	788	TGACTATTGTGATAGCATCAAGGTTATCGGGGGCAACAGCTCTGAACCTTATTCGTCA	847				
QY	843	ACATGGGTCATAGAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGA	902				

RESULT 2
 BG837708
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

BG837708
 Zm10_01f08_A Zm10_AAFCEC0RC_Fusarium_graminearum_corn_silk Zea
 mays cDNA clone Zm10_01f08, mRNA sequence.
 BG837708
 BG837708.1 GI:14204031
 Zea mays.
 Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 901)
 Morris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
 Haars,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A, Sprott
 D. and Tinker,N.A.
 Expressed Sequence Tags from Maize Silk Six Hours After Silk
 Channel Inoculation with Fusarium graminearum
 Unpublished (2001)
 Contact: Harris, Linda J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1314
 Fax: (613) 759-6566
 Email: harrislj@em.agr.ca.
 Location/Qualifiers
 1..901
 /organism="Zea mays"
 /cultivar="CO388"
 /db_xref="taxon:4577"
 /clone="Zm10_01f08"

```
/clone_lib="zm10_AAFc_ECORC_Fusarium_graminearum_corn_silk"
/tissue_type="Silk"
/dev_stage="4-5 days post-silk emergence"
/site="XhoI; EcoRI; Site_1: EcoRI; Site_2: XhoI; Field-grown corn was silk channel-inoculated in the morning (~10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of Fusarium graminearum and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel. RNA from five silk channels was pooled."

BASE COUNT      267 a      208 g      249 t      3 others
ORIGIN

Query Match      53.2%; Score 778.8; DB 12; Length 901;
Best Local Similarity 98.3%; Pred. No. 3.8e-133;
Matches 816; Conservative 2; Mismatches 9; Indels 3; Gaps 3;

QY 605 TCGCTGTTGCTTCAGAGATATGACCTCCCTTACTTTTGGGCTCCACGGTTCCTTCGTC 664
DB 1 TCGCTGTTGCTTCAGAGATATGACCTCCCTTACTTTTGGGCTCCACGGTTCCTTCGTC 60

QY 665 ATTTAATGATCCCAAGTTCACAGAAATACCTGTGATGGAATTTGATGTTGCCAAGTTT 724
DB 61 ATTTAATGGAYCAAAGTTCCACAGAAATACCTGTGATGGAATTTGATGTTGCCAAGTTT 120

QY 725 TGGAGGAGCTTGACCTACCATGACCATGACATGATGATTTGTGATCCTGTGTGGATGTG 784
DB 121 TGGAGGAGCTTGACCTACCATGACCATGACATGATGATTTGTGATCCTGTGTGGATGTG 180

QY 785 ACTATTCTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAACTTATTCGTCAAC 844
DB 181 ACTATTCTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAACTTATTCGTCAAC 240

QY 845 ATGGGTCATAGAAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATCC-TGAG 903
DB 241 ATGGGTCATAGAAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATCCWTGAG 300

QY 904 GACTGGCTTACC-AAGAACTCGACCTTGTTCAAAGAGCCATATGCATTTGGATAT 962
DB 301 GACTGGCTTACCAGAAAGCTGACCTTGTTCAAAGAGCCATATGCATTTGGATAT 360

QY 963 TCTTGAGCTAAATGGAAGTCACTGATGAGGAGGCTCTATAGTTTCTTGTTAAAGA 1022
DB 361 TCTTGAGCTAAATGGAAGTCACTGATGAGGAGGCTCTATAGTTTCTTGTTAAAGA 420

QY 1023 TAATGGTTTCAACGAAGATCGGGTGAACAAGGCCATAGAGAAGATCAAAATCTGCCAGAA 1082
DB 421 TAATGGTTTCAATGAAGATCGGGTGAACAAGGCCATAGAGAAGATCAAAATCTGCCAGAA 480

QY 1083 TAATCTCCGAGGAGACTCGAGTCTTTTCAAGCCACTCCACCATCAGCACC 1142
DB 481 TAATCTCCGAGGAGACTCGAGTCTTTTCAAGCCACTCCACCATCAGCACC 540

QY 1143 GCTAAAAACGGAAGAGACTTCGGATATAAACAAGCAGCTCGCAACAGAAACAAA 1202
DB 541 GCTAAACGGAAGAGACTTCGGATATAAACAAGCAGCAGCTCGCAACAGAAACAAA 600

QY 1203 GCGTGGTGAAGAAATAATCTTGATGCTTTGATGTACAACTACGACTACGAAAGCA 1262
DB 601 GCGTGGTGAAGAAATAATCTTGATGCTTTGATGTACAACTACGACTACGAAAGCA 660

QY 1263 GCGTGGGTGATCAGTTCGTTAGATATTTAACTCCGTGTTTAACTCAGACCTTTGG 1322
DB 661 GCGTGGGTGATCAGTTCGTTAGATATTTAACTCCGTGTTTAACTCAGACCTTTGG 720

QY 1323 TAAAGTTTGTCTATGTTTCAAGCTGGGTAAAGTTAGTTGTTGTTGAAGAGATTGGTGA 1382
DB 721 TGAAGTTTGTCCCAATGTTTCAAGCTGGGTAAAGTTAGTTGTTGTTGAAGAGATTGGTGA 780

QY 1383 CCAAGTAACAAACTTATCGCTG-TTTTTTACTTCTTGTCTTCTTGAAGTA 1431
```

```
||||| 781 CCAAGTAACAAACTTATCGCTGTTTTTTTACTTCTTGTCTTGAAGTA 830
||||| 781 CCAAGTAACAAACTTATCGCTGTTTTTTTACTTCTTGTCTTGAAGTA 830

RESULT 3
BU037559 638 bp mRNA linear EST 23-AUG-2002
LOCUS 946140C05.v1 946 - tassell primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BU037559
VERSION BU037559.1 GI:22473079
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 638)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946140 row: C column: 05.
Location/Qualifiers
FEATURES
source
1..638
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassell primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XL0LR"
/note="Organ: tassels; Vector: HybrizAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybrizAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT 196 a 141 c 171 g 130 t
ORIGIN

Query Match 42.0%; Score 614.8; DB 14; Length 638;
Best Local Similarity 98.7%; Pred. No. 4.9e-103;
Matches 630; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 2 ACAGAAATAGTCTCGCGCGGGTTTCTTGGCCACTCCGCTCAGCGCGCGCGCCAC 61
DB 2 ACAGAAATAGTCTCGCGTTCGGTTTCTTGGCCACTCCGCTCAGCGCGCGCGCCAC 61

QY 62 CCGCCACAGCGCGCGAGAGATGGCATCAAGGTTTCACGAAATCTGCGGACA 121
DB 62 CCGCCACAGCGCGCGAGAGATGGCATC-AGGTTTTCACGAAATCTGCGGACA 120

QY 122 ATGCGCCCAAGCGGATGAAGGAGCAGAGTTTCGAGAGCTACTTGGCGCGCAATTCGCG 181
DB 121 ATGCGCCCAAGCGGATGAAGGAGCAGAGTTTCGAGAGCTACTTGGCGCGCAATTCGCG 180

QY 182 TCGAGCGCAGCATGAGCATATACAGTTCTGATGTAGTTGGAGAGACAGCATGGAAA 241
DB 181 TCGAGCGCAGCATGAGCATATACAGTTCTCTGATAGTAGTTGGAGAGACAGCATGGAAA 240

QY 242 CTCTCACAAATGAAGCTGGTGAAGTCACTACTACTATTTTCAAGGATGTTCAACCGACAA 301
DB 241 CTCTCACAAATGAAGCTGGTGAAGTCACTACTACTATTTTCAAGGATGTTCAACCGACAA 300
```

```

QY 302 TAAGATTACTGGAAGCGGGAATCAAGCCAGCTTTATGTTTTTGTATGGCAAGCCTCTCTGATA 361
Db 301 TAAGATTACTGGAAGCGGGAATCAAGCCAGCTTTATGTTTTTGTATGGCAAGCCTCTCTGATA 360
QY 362 TGAAGAAAACAAGAGCTTCTCTAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGA 421
Db 361 TGAAGAAACAGACTTCTCTAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGA 420
QY 422 CTGAGGCAGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAA 481
Db 421 CTGAGGCAGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAA 480
QY 482 AGGTCACAAGGCACACACAACGAAGATTGTAACGCTATTAAAGACTTATGAGCAAGAGGACTGTAA 541
Db 481 AGGTCACAAGGCACACACAACGAAGATTGTAACGCTATTAAAGACTTATGAGCAAGAGGACTGTAA 540
QY 542 TTGTAGAGGCACCTTCTCAAGCAGAGCAGAAATGTCAGCGCCTTTGCGAGCCTTTGCGATAAACGATAAGG 601
Db 541 TTGTAGAGGCACCTTCTCAAGCAGAGCAGAAATGTCAGCGCCTTTGCGAGCCTTTGCGATAAACGATAAGG 600
QY 602 TGTTCGCTGTGCTCTCAAGATATGGACTCCCTTACT 639
Db 601 TGTTCGCTGTGCTCTCAAGATATGGACTCCCTTACT 638

RESULT 4
A1881599 586 bp mRNA linear EST 02-FEB-2000
LOCUS 606068G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION A1881599
VERSION A1881599.1 GI:5566733
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
Clade: Panicoideae; Andropogoneae; Zea.
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.
Location/Qualifiers
1. 586
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt
lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOR (Stratagene)"
/notes="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"

BASE COUNT 186 a 132 c 159 g 109 t
ORIGIN
Query Match 38.38; Score 560.8; DB 9; Length 586;
Best Local Similarity 98.88; Pred. No. 3.9e-93;
Matches 563; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ACGAAGATAGTCGCGCGCGGTTCTTTCGCGCACTCCGGCTCAGCGCGCGCGCCGAC 61
||| ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||

```

```

Db 15 ACGGAATAGTCGCGGTCGCGGTTTCTTTCGCGCACTCCGGCTCAGCGCGCGCGCCAC 74
QY 62 CCGCCACAGCCGCGCAGACGAGATGGCATCAAGGGTTTTCACGAAACTGCTGGTGGACA 121
Db 75 CCGCCACAGCCGCGCAGACGAGATGGCATCAAGGGTTTTCACGAAACTGCTGGTGGACA 134
QY 122 ATGGCCCAAGGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGG 181
Db 135 ATGGCCCAAGGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGG 194
QY 182 TCGAGCCAGCATGAGCATATACCAGTTCCTTGATTGTAGTTGGAGGAGCAGCATGGAAA 241
Db 195 TCGAGCCAGCATGAGCATATACCAGTTCCTTGATTGTAGTTGGAGGAGCAGCATGGAAA 254
QY 242 CTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTTCAAGGAATGTTCAACCGGACAA 301
Db 255 CTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTTCAAGGAATGTTCAACCGGACAA 314
QY 302 TAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTGTATGGCAAGCCTCTCTGATA 361
Db 315 TAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTGTATGGCAAGCCTCTCTGATA 374
QY 362 TGAAGAAAACAAGAGCTTCTCTAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGA 421
Db 375 TGAAGAAAACAAGAGCTTCTCTAAAGATACTCAAAAAGAGATGATGCAACCAAGATCTGA 434
QY 422 CTGAGGCAGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAA 481
Db 435 CTGAGGCAGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAA 494
QY 482 AGGTCACAAGGCACACACAACGAAGATTGTAACGCTATTAAAGACTTATGAGCAAGAGGACTGT 541
Db 495 AGGTCACAAGGCACACACAACGAAGATTGTAACGCTATTAAAGACTTATGAGCAAGAGGACTGT 554
QY 542 TTGTAGAGGCACCTTCTCAAGCAGAGCAGAA 573
Db 555 TTGTAGAGGCACCTTCTCAAGCAGAGCAGAA 586

RESULT 5
BE639422 550 bp mRNA linear FST 30-AUG-2000
LOCUS 946033A02.y2 946 - tassell primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BE639422
VERSION BE639422.1 GI:9952839
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
Clade: Panicoideae; Andropogoneae; Zea.
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
Location/Qualifiers
1. 550
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassell primordium prepared by Schmidt
lab"
/tissue_type="tassels"

FEATURES
source

```

/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybridZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."
BASE COUNT 177 a 116 c 149 g 108 t
ORIGIN

Query Match 36.7%; Score 536.4; DB 10; Length 550;
Best Local Similarity 99.6%; Pred. No. 1.2e-88;
Matches 548; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 46 AGCGCGCGCGCCACCGCCGCGGCGGAGAGAGATGGGCAATGAGGTTTGACG 105
Db 1 AGCGCGCGCGCGCCACCGCCGCGGCGGAGAGAGATGGGCAATGAGGTTTGACG 60
QY 106 AAATGCTGCGGAGCAATGCGCCCAAGGGGATGAAGGAGCAGAAAGTTCGAGAGCTACTTC 165
Db 61 AAATGCTGCGGAGCAATGCGCCCAAGGGGATGAAGGAGCAGAAAGTTCGAGAGCTACTTC 120
QY 166 GCGCGCAAAATCCGTCGACGCGCAGCATATACCATGTCCTGATGTTAGTTGGA 225
Db 121 GCGCGCAAAATCCGTCGACGCGCAGCATATACCATGTCCTGATGTTAGTTGGA 180
QY 226 AGGACAGG-CATGGAACTCTCAAAATGAAGTGGTGAAGTCACTAGTCAATTTGCAAGG 284
Db 181 AGGACAGGTCATGGAACTCTCAAAATGAAGTGGTGAAGTCACTAGTCAATTTGCAAGG 240
QY 285 AATGTTCAACCGGACAAATAGATTACTTGAAGCGGGAATCAAGCCAGTTTATGTTTTGA 344
Db 241 AATGTTCAACCGGACAAATAGATTACTTGAAGCGGGAATCAAGCCAGTTTATGTTTTGA 300
QY 345 TGGCAAGCTTCCTGATATGAAGAAACAAGAGCTTGCTGATAAGATCTCAAAAAGAGATGA 404
Db 301 TGGCAAGCTTCCTGATATGAAGAAACAAGAGCTTGCTGATAAGATCTCAAAAAGAGATGA 360
QY 405 TGCACCAAGAGATCTGACTGAGCAGCTAGAGTAGGATGAAGATGAAGTTCGAAATTT 464
Db 361 TGCACCAAGAGATCTGACTGAGCAGCTAGAGTAGGATGAAGATGAAGTTCGAAATTT 420
QY 465 GAGCAAGAGCTGTAAAGTCTCAAGGCAACACAGCAAGATTTGTAAGCGGTATTAAAG 524
Db 421 GAGCAAGAGCTGTAAAGTCTCAAGGCAACACAGCAAGATTTGTAAGCGGTATTAAAG 480
QY 525 ACTTATGGGGTTCCTGTTAGAGGACCTTCTGAAAGCAGAGCAGAAATGTGCGAGCCT 584
Db 481 ACTTATGGGGTTCCTGTTAGAGGACCTTCTGAAAGCAGAGCAGAAATGTGCGAGCCT 540
QY 585 TTGCATAAAC 594
Db 541 TTGCATAAAC 550

RESULT 6
BE639421/c 554 bp mRNA linear EST 30-AUG-2000
LOCUS
DEFINITION 946033A02.x2 946 - tassel primordium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.

ACCESSION BE639421
VERSION BE639421.1 GI:9952838
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 554)
AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.

FEATURES
source Location/Qualifiers
1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassel primordium prepared by Schmidt lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybridZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."
BASE COUNT 145 a 131 c 110 g 168 t
ORIGIN

Query Match 36.5%; Score 534.4; DB 10; Length 554;
Best Local Similarity 99.8%; Pred. No. 2.8e-88;
Matches 535; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 896 TTCTGAGGACTGGCTTACCAAGAGCTCGAGCTTTGTTCAAGGAGCCCTAATGTTCACAT 955
Db 554 TTCTGAGGACTGGCTTACCAAGAGCTCGAGCTTTGTTCAAGGAGCCCTAATGTTCACAT 495
QY 956 TGGATATTCCTGAGCTAAATGAGACTGCACCTGATGAGGAGGCTCATTAAGTTTCTCTGG 1015
Db 494 TGGATATTCCTGAGCTAAATGAGCTGCACCTGATGAGGAGGCTCATTAAGTTTCTCTGG 435
QY 1016 TAAAGATATGTTTCAACGAGAGATCGGTGACAAAGGCATAGAGAAGATCAATCTG 1075
Db 434 TAAAGATATGTTTCAACGAGAGATCGGTGACAAAGGCATAGAGAAGATCAATCTG 375
QY 1076 CCAAGATAAATCGTCGCAAGGAGACTCGAGTCTTTTCAAGCCCAACTGCCACCAT 1135
Db 374 CCAAGATAAATCGTCGCAAGGAGACTCGAGTCTTTTCAAGCCCAACTGCCACCAT 315
QY 1136 CAGCACCGCTAAACGGAAGAGACTTCGGATATAAACAAGCAAGGAGCTCCGAAACAGA 1195
Db 314 CAGCACCGCTAAACGGAAGAGACTTCGGATATAAACAAGCAAGGAGCTCCGAAACAGA 255
QY 1196 AACAAAGGCTGGTGGAAAGAGAAATATCTCGATGCTTGTACACTAGACTAC 1255
Db 254 AACAAAGGCTGGTGGAAAGAGAAATATCTCGATGCTTGTACACTAGACTAC 195
QY 1256 GAAAGCAGCGTGGCTGATCAGTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGA 1315
Db 194 GAAAGCAGCGTGGCTGATCAGTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGA 135
QY 1316 GCTTTGGTAAAGTTTGCTCATGTTTCAAGCTGGGGTAAGTAGTTGTTGTTGAAGAT 1375
Db 134 GCTTTGGTAAAGTTTGCTCATGTTTCAAGCTGGGGTAAGTAGTTGTTGTTGAAGAT 75
QY 1376 TGGTGTACCAAGTAAACAACTTATCGCTGTTTTTACTTCTTCTGCTTTCGAAGTA 1431
Db 74 TGGTGTACCAAGTAAACAACTTATCGCTGTTTTTACTTCTTCTGCTTTCGAAGTA 19
RESULT 7
AW562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS
DEFINITION 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,

```

mRNA sequence.
ACCESSION AW562789.1 GI:7216667
VERSION AW562789.1
KEYWORDS EST.
SOURCE zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 60065 row: H column: 06.
FEATURES
source
1. .553
Location/Qualifiers
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 171 a 111 c 129 g 142 t
ORIGIN
Query Match 35.4%; Score 517.8; DB 10; Length 553;
Best Local Similarity 98.7%; Pred. No. 3e-85;
Matches 522; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 903 GGACTGGCCTTACCAAGAGCTCGACGCTTGTTCAGGAGGCTATGTCACATGGATAT 962
Db 4 GGACTGGCCTTACCAAGAGCTCGACGCTTGTTCAGGAGGCTATGTCACATGGATAT 63
Qy 963 TCCTGAGCTAAATGGATCGACCTGATGAGGAGGCTCTATAAGTTTCCTGGTAAAGA 1022
Db 64 TCCTGAGCTAAATGGATCGACCTGATGAGGAGGCTCTATAAGTTTCCTGGTAAAGA 123
Qy 1023 TAATGGTTTCAACGAAGATCGGTTGACAAAGCCATAGAGATCAATCTGCCAAGA 1082
Db 124 TAATGGTTTCAATGAGATCGGGTGACAAAGCCATAGAGATCAATCTGCCAAGA 183
Qy 1083 TAATGCTGCGAAGAGACTCGAGTCTCTTTTCAAGCCAACTGCCACATCAGACCC 1142
Db 184 TAATGCTGCGAAGAGACTCGAGTCTCTTTTCAAGCCAACTGCCACATCAGACCC 243
Qy 1143 GCTAAACGGAGAGACTTCGGTATTAACGAAGCAGCCTCGGATACAGAAACAAA 1202
Db 244 GCTAAACGGAGAGACTTCGGTATTAACGAAGCAGCCTCGGATACAGAAACAAA 303
Qy 1203 GGCTGGTGGAAAGAGAAATATCTTGATGCTTGTATGATCAACTACGACTACGAAGCA 1262
Db 304 GGCTGGTGGAAAGAGAAATATCTTGATGCTTGTATGATCAACTACGACTACGAAGCA 363
Qy 1263 GCGGTGGCGTATCAGCTTCGCTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTGG 1322
Db 364 GCGGTGGCATGATCAGCTTCGCTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTGG 423
Qy 1323 TAAAGTTTCTCATCTTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATGGTGTA 1382
Db 424 TGAAGTTTGCCTCATCTTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATGGTGTA 483

```

```

Qy 1383 CCAAGTACAAACATATCGCTGTTTTTTTACTTCTTGTCTCTTTGAAGTA 1431
Db 484 CCAAGTACAAACATATCGCTGCTTTTACTTCTTGTCTCTTTGAAGTA 532
RESULT 8
AW000375 532 bp mRNA linear EST 08-SEP-1999
LOCUS 614014D03.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW000375
VERSION AW000375.1 GI:5847296
KEYWORDS EST.
SOURCE zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614014 row: D column: 03.
FEATURES
source
1. .532
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOR"
/note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
BASE COUNT 173 a 103 c 129 g 127 t
ORIGIN
Query Match 34.2%; Score 501; DB 10; Length 532;
Best Local Similarity 98.7%; Pred. No. 3.7e-82;
Matches 526; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
Qy 666 TTTAATGATCCAAAGTTCCAAAGAAATACCTGTGATGGAATTTGATGTGCCAAGTTT 725
Db 1 TTTAATGATCCAAAGTTTCAAGAAATACCTGTGATGGAATTTGATGTGCCAAGTTT 60
Qy 726 GGAGAGCTTCAACTCACCATGGACCATTCATTTGATTTGTCATCTCTGTGGATGTA 785
Db 61 GGAGAGCTTCAACTCACCATGGACCATTCATTTGATTTGTCATCTCTGTGGATGTA 120
Qy 786 CTATTGTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTTCGTAACA 845
Db 121 CTATTGTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTTCGTAACA 180
Qy 846 TGGTTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACA -GATATCAAAATTCCTGAG 904
Db 181 TGGTTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAAGGATATCAAAATTCCTGAG 240
Qy 905 ACTGGCTTACCAGAGCTCGACGCTTGTTCAGGAGCCCTAATGTCACATGGATATTC 964
Db 241 ACTGGCTTACCAGAGCTCGACGCTTGTTCAGGAGCCCTAATGTC -CATTTGATATTC 299
Qy 965 CTGAGCTAAAATGACCTGCACCTGATGAGGAGGTCTCATAGTTTCTCGTAAAAAGATA 1024

```

Qy	275	ATTTGCAAGGAATGTTCAACCGGACAATAAGATTACTGGGAAGCGGGAATCAAGCCAGTTT	334
Db	272	ATTTGCAAGGCATGTTTCAGCCGGACAATAAGGTTGCTCGAGGCAGGAATTAACCCAGTAT	331
Qy	335	ATGTTTTTGATGGCAAGCCTCTCATATGAAGAACAAGAGCTTGCTAAAAGATACTCAAA	394
Db	332	ATGTTTTTGATGGCAAGCCTCTCTGAAATGAAGAAGGACGAGCTTTTAAAAAGACACGCA	391
Qy	395	AAAGAGATGATGTCACCAACCAAGATCTGACTGTAGGCGAGTAGAGGTAGGAGATAAAGATGCGA	454
Db	392	AGAGAAATGAAGCAACAGAGAGCTGACGAAGCGCATAGAGCGAGGATACGGATGCAA	451
Qy	455	TTGAAAATTTGAGCAAGAGGACTGTAAAGGTCACAGGCAACACACAGCAAGATTTGTAAC	514
Db	452	TTGAAAATTTGAGCAAGAGGACTGTAAAGGTCACAGGCAACACACAGCAAGATTTGTAAC	511
Qy	515	GGCTATTAAAGACTTATGGGGGTTCTGTTGTAGAGGCACTTCTGAAGCAAAACCAAGAT	574
Db	512	GTCTACTAAAGACTGATGGGTTCTGTTGTAGAGGCTCTTGTGAAGCAAGATCAACAT	571
Qy	575	GTGACGCCCTTTGATAAACGATAAGGTTTCGGCTTTCAGAAAGATTTGGAATCC	634
Db	572	GTGTCGCCCTTTGACAGAGTGACAAAGGTGTATGCTGTTGCATCAGAAAGATTTGGAATCC	631
Qy	635	TTACTTTTGGGGCTCCA	651
Db	632	TTACTTTTGGAGTCCA	648

RESULT 10
 AW559173/c
 LOCUS
 DEFINITION 660065H06.x1 660 - Mixed stages of anther and pollen Zea mays cdna. 474 bp mRNA linear EST 07-MAR-2000

ACCESSION AW559173
 VERSION AW559173.1 GI:7204640
 KEYWORDS EST.
 SOURCE zeamays
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 474)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cdna libraries sequenced at Stanford University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers
 1. 474
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI; Site_2: XhoI; Anther and pollen cdna library. Directionally sequenced with 5' end at the EcoRI site. Created by Amie Franklin."

BASE COUNT 113 a 114 c 92 g 154 t 1 others
 ORIGIN

Query Match 31.7%; Score 453.4; DB 10; Length 474;
 Best Local Similarity 98.5%; Pred No. 3e-75;

Matches 467; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 920 AAGTCGACGCTGTTCAGGAGCCTAATGTACATTTGATATTCCTGAGCTAAATGCA 979
 DB 474 AAGTCGACGCTGTTCAGGAGCCTAATGTACATTTGATATTCCTGAGCTAAATGCA 415
 QY 980 CTGCACCTGATGAGGAGGCTCTCATTAAGTTTCTGTTAAAGATTAATGTTTCAACGAAG 1039
 DB 414 CTGCACCTGATGAGGAGGCTCTCATTAAGTTTCTGTTAAAGATTAATGTTTCAACGAAG 355
 QY 1040 ATCCGGTGCACAAAGCCATAGAGAATCAATCTGCCAAGATAATTCGTCGCAAGAA 1099
 DB 354 ATCCGGTGCACAAAGCCATAGAGAATCAATCTGCCAAGATAATTCGTCGCAAGAA 295
 QY 1100 GACTCGAGTCCTTTTCAAGCCAACTGCCACCATCAGCACCGCTAAACCGGAAGGAGA 1159
 DB 294 GACTCGAGTCCTTTTCAAGCCAACTGCCACCATCAGCACCGCTAAACCGGAAGGAGA 235
 QY 1160 CTTCGGATAAACAAGCAAGGAGCTGCGAACAGAAACAAGGCTGGTGGAAAGAGA 1219
 DB 234 CTTCGGATAAACAAGCAAGGAGCTGCGAACAGAAACAAGGCTGGTGGAAAGAGA 175
 QY 1220 AATAATCTTGGATGCTGATGTACAACTACGACTACGAAAGCAGCGTGGCGTGATCACT 1279
 DB 174 AATAATCTTGGATGCTGATGTACAACTACGACTACGAAAGCAGCGTGGCGTGATCACT 115
 QY 1280 TCGCTTAGATTTAATCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGTCTCATGT 1339
 DB 114 TCGCTTAGATTTAATCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGTCCATGT 55
 QY 1340 TTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGGTGACCAAGTAACAA 1393
 DB 54 TTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGGTGACCAAGTAACAA 1

RESULT 11
 AI861468/c
 LOCUS 614014D03.xl 614 - root cDNA library from Walbot Lab Zea mays cDNA, EST 19-JUL-1999
 DEFINITION mRNA sequence.
 ACCESSION AI861468
 VERSION AI861468.1 GI:5525575
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 Zea mays 1 (bases 1 to 470)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford University
 Unpublished (1999)
 CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 614014 row: D column: 03.
 Location/Qualifiers
 1. 470
 /organism="Zea mays"
 /cultivar="W23"
 /db_xref="taxon:4577"
 /clone_lib="614 - root cDNA library from Walbot Lab"
 /tissue_type="root"
 /dev_stage="3-4 days old"
 /lab_host="X10LR"
 /note="Organ: root; Vector: pBluescriptII SK+; Site: 1: EcoRI; Site: 2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"

BASE COUNT 112 a 113 c 93 g 152 t

ORIGIN

Query Match 31.6%; Score 462; DB 9; Length 470;
 Best Local Similarity 98.9%; Pred. No. 5.4e-75;
 Matches 465; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 937 AAGGAGCCTAATGTACATTTGATATTCCTGAGCTAAATGACATGCACCTGATGAGAG 996
 DB 470 AAGGAGCCTAATGTCCCATTTGATATTCCTGAGCTAAATGACATGCACCTGATGAGAG 411
 QY 997 GGTCTCATAAGTTTCTGTTAAAGATTAATGTTTCAACGAAGATCGGTGACAAAGGCC 1056
 DB 410 GGTCTCATAAGTTTCTGTTAAAGATTAATGTTTCAACGAAGATCGGTGACAAAGGCC 351
 QY 1057 ATAGAGAAGATCAATCTGCCAAGATAATTCGTCGCAAGAGACTCGAGTCTCTTTTC 1116
 DB 350 ATAGAGAAGATCAATCTGCCAAGATAATTCGTCGCAAGAGACTCGAGTCTCTTTTC 291
 QY 1117 AAGCCAACTGCCACCATCAGCACCGCTAAACCGGAAGGAGCTTCGGATAAAACAAGC 1176
 DB 290 AAGCCAACTGCCACCATCAGCACCGCTAAACCGGAAGGAGCTTCGGATAAAACAAGC 231
 QY 1177 AAGGAGCTGCGAAACAAGAAACAAGGCTGGTGGAAAGAAATAATCTTGGATGCTT 1236
 DB 230 AAGGAGCTGCGAAACAAGAAACAAGGCTGGTGGAAAGAAATAATCTTGGATGCTT 171
 QY 1237 GATGTACAACTACGACTACGAAAGCAGCGTGGCGTGATCACTTCGCTTAGATTTAA 1296
 DB 170 GATGTACAACTACGACTACGAAAGCAGCGTGGCGTGATCACTTCGCTTAGATTTAA 111
 QY 1297 CTCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGCCTCATGTTTCAAGCTGGGGTAAAGT 1356
 DB 110 CTCCTGTTTAACTCAGAGCTTTGGTAAAGTTTGCCTCATGTTTCAAGCTGGGGTAAAGT 51
 QY 1357 TAGTTGTTTGAAGAGATTGGTGACCAAGTAAACAACTTATCGCTGT 1406
 DB 50 TAGTTGTTTGAAGAGATTGGTGACCAAGTAAACAACTTATCGCTGT 1

RESULT 12
 BQ986894
 LOCUS BQ986894
 DEFINITION BQ986894.1 yg.abl QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
 ACCESSION BQ986894
 VERSION BQ986894.1 GI:22404419
 KEYWORDS EST.
 SOURCE Lactuca sativa.
 ORGANISM Lactuca sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.
 Lactuca 1 (bases 1 to 735)
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compositae.ucdavis.edu/
 Unpublished (2002)
 CONTACT: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
 singleton, see http://cgdb.ucdavis.edu/ for details.
 Plate: QGF10 row: L column: 19.
 Location/Qualifiers
 1. 735
 source


```
/organism="Lactuca sativa"  
/cultivar="L. serriola"  
/db_xref="taxon:4236"  
/clone_lib="QG_EFGHJ lettuce serriola"  
/lab_host="E.coli"  
/note="Vector: pBRCDNASFIAB; The library was constructed  
from 10 different sources of RNA from a single genotype.  
Separate cDNAs were generated using primers that  
incorporated unique 5' and 3' tags to distinguish each  
source of RNA. cDNAs were then pooled, size-fractionated,  
directionally cloned into a custom medium-copy vector and  
transformations made with four size classes to minimize  
size bias. Details of each source of RNA and library  
construction can be obtained at http://cgdb.ucdavis.edu/  
TAG_LIB=QG_EFGHJ lettuce serriola  
TAG_TISSUE=flowers post-fertilized  
TAG_SEQ=TCGCATCGGG"  
BASE COUNT 227 a 121 c 178 g 207 t 2 others  
ORIGIN  
Query Match 31.2%; Score 456.4; DB 14; Length 735;  
Best Local Similarity 77.7%; Pred. No. 4.9e-74;  
Matches 550; Conservative 0; Mismatches 158; Indels 0; Gaps 0;  
QY 232 GGCATGGAACCTCTCAAAATGAAGCTGGTGAAGTCACTAGTATTCGAAAGGAATGTC 291  
Db 1 GGGACAGAAATGCTGACCAATGAGGTGGTGTGCTGCTGCTAGTCCATTTGCAAGGAATGTC 60  
QY 292 AACCGGACATAGATTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTGTGATGCGAAG 351  
Db 61 AATCGGACTATTAGGCTTCTTGAATCTGGATTGAAGCCAGTCTATGTTTGTGATGCGGCA 120  
QY 352 CCTCCTGATATGAAGAACAGAGCTTGTAAAGATACCTCAAAAGAGATGATGCAACC 411  
Db 121 CCTCCGATTGAAGAACAGAACTTGCAAAAGATACCTCAAGCGGAGAGATGCAACA 180  
QY 412 AAGATCTGACTGAGGCAGTAGAGTGGAGATGAAGATCGGATGAAATTTGAGCAAG 471  
Db 181 GCGGAGTTAGATGAGGCTATTAACTGCGCAATAAGGAAGACATTTGAAATTTAGTAA 240  
QY 472 AGGACTTAAAGTCAAGGCAACACAGCAAGATTTGAAGCGCTATTAAAGCTTTATG 531  
Db 241 CGTATGTGAAGTAAACAGACACAATGAGGATTTGCAAAAACCTCTTAAGGCTTATG 300  
QY 532 GGGGTTCTCTGTAGAGGCACCTTCTGAAGCAGAGCAGCAATGTGAGCCCTTTGCATA 591  
Db 301 GGAGTGCCTGTTATTGAGGCCCATCTGAAGCAGAGGCACAATGTGCTCTTTGCAAA 360  
QY 592 AACGATAAGGTGTTGCTGTGCTTCAGAAAGATATGGACTCCCTTACTTTTGGGCTCCA 651  
Db 361 GCTGATAAGGTGTATGCTGTGCTCTGAAGATATGGATTCTCTACTTTTGGAGCACCA 420  
QY 652 CGGTTCTTCTGTCATTTAATGATCCCAAGTTTCCAAGAAATACCTGTATGGAATTGAT 711  
Db 421 AATTTCTTAGACATTTAATGGATTCCTGCTTCAAGAAATATCCCGTATGGAATTGAT 480  
QY 712 GTTGCCAAAGTTTGGAGGAGCTTGAACCTCACCATGGACAGCTTCAATTTGTTGTCATC 771  
Db 481 GTTTCCAAAGTTTGAAGGAACCTGAACCTTACATTTAGATCAATTTATGATCTGTATC 540  
QY 772 CTGTGTGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGGCAACAGCTCTGAAA 831  
Db 541 CTATGTGGATGTGATTATTGTGACGATATTAGAGGTATTGTTGGCGGACAGCTTTGAAG 600  
QY 832 CTTATTTCGTCACATGGTCCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGATAT 891  
Db 601 CTCATTTCGTCACATGGTCCATAGANACTATCTCTGAAAATATAAACAAGACAGATAT 660  
QY 892 CAAATTCCTGAGGACTGGCCCTTACCAGAAAGCTCGAGCTTGTTCAG 939  
Db 661 CANATACCTGAGAAATGGCCATATCAGGAGGCTAGACGCCCTTTTCAG 708
```

RESULT 13
BE186786/c

```
LOCUS  
DEFINITION 467 bp mRNA linear EST 22-JUN-2000  
946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea  
mays CDNA, mRNA sequence.  
ACCESSION BE186786  
VERSION BE186786.1 GI:8665970  
KEYWORDS EST.  
SOURCE Zea mays.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 467)  
Walbot, V.  
Maize ESTs from various cDNA libraries sequenced at Stanford  
University  
Contact: Walbot V  
Unpublished (1999)  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 946012 row: C column: 08.  
Location/Qualifiers  
1. 467  
/organism="Zea mays"  
/cultivar="OH43"  
/db_xref="taxon:4577"  
/clone_lib="946 - tassal primordium prepared by Schmidt  
lab"  
/tissue_type="tassels"  
/dev_stage="just after the transition from vegetative to  
inflorescence development"  
/lab_host="XL0LR"  
/note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;  
Site_2: XhoI; George Chuck dissected immature tassels  
between Imm and 3mm. Sharon Stanfield prepared the cDNA  
library in HybriZAP. Sample insert size range was 350 bp  
to 3 Kb with a 1 Kb average."  
BASE COUNT 108 a 110 c 96 g 153 t  
ORIGIN  
Query Match 30.6%; Score 448; DB 10; Length 467;  
Best Local Similarity 98.7%; Pred. No. 2e-72;  
Matches 462; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 886 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAGAAGCTCGACCTGTTCAGGAGCCT 945  
Db 467 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAGAAGCTCGACCTGTTCAGGAGCCT 408  
QY 946 AATGTCAATTTGGATATTCCTGAGCTAAATGGACTGCACCTGTATGAGGAGGTCTCATA 1005  
Db 407 AATGTCAATTTGGATATTCCTGAGCTAAATGGACTGCACCTGTATGAGGAGGTCTCATA 348  
QY 1006 AGTTTCTGTGTAAGAATATGTTTCAACGAAGATCGGTGACAAAGGCCATAGAGAAG 1065  
Db 347 AGTTTCTGTGTAAGAATATGTTTCAATGAAGATCGGTGACAAAGGCCATAGAGAAG 288  
QY 1066 ATCAATATCTCCCAAGAAATAATCGTCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACT 1125  
Db 287 ATCAATATCTCCCAAGAAATAATCGTCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACT 228  
QY 1126 GCCACCACATCAGCACCGCTTAAACGGAAGAGACTTCGGATATAAACAAGCAGGAGCT 1185  
Db 227 GCCACCACATCAGCACCGCTTAAACGGAAGAGACTTCGGATATAAACAAGCAGGAGCT 168  
QY 1186 GCCAACAAGAAAACAAGAGCTGGTGGAAAGAAATAATCTTGGATGCTTGTATGACAA 1245  
Db 167 GCCAACAAGAAAACAAGAGCTGGTGGAAAGAAATAATCTTGGATGCTTGTATGACAA 108
```



```

QY 1246 CTAGACTACGAAAGCAGCGGTGGCGTGATCACATTCGCTTAGATATATTAACTCCCGTT 1305
|||||
Db 107 CTAGACTACGAAAGCAGCGGTGGCGATGATCACATTCGCTAGATATATTAACTCCCGTT 49
|||||
QY 1306 TTAAGTCTAGAGCTTTGTGAAAGTTTCTCTATGTTTCAAGCTGGGTA 1353
|||||
Db 48 TTAAGTCTAGAGCTTTGTGAAAGTTTCTCTATGTTTCAAGCTGGGTA 1
|||||

RESULT 14
AW562517/c 475 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.X2 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW562517.1 GI:7216395
VERSION AW562517
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE clade; Panicoideae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 475)
COMMENT Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
Location/Qualifiers
1. 475
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 124 a 110 c 94 g 147 t
ORIGIN
Query Match 29.4%; Score 429.8; DB 10; Length 475;
Best Local Similarity 98.2%; Pred. No. 4.3e-69;
Matches 445; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 979 ACTGCACCTGATGAGGAGGTTCTATAGTTTCTGTTTAAAGATAATATGTTTCAACGAA 1038
|||||
Db 475 ACTGCACCTGATGAGGAGGTTCTATAGTTTCTGTTTAAAGATAATATGTTTCAATGAA 416
|||||
QY 1039 GATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAATATCGTCGCAAGGA 1098
|||||
Db 415 GATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAATATCGTCGCAAGGA 356
|||||
QY 1099 AGACTCGAGTCTTTTCAAGCCAACTGCCACCATCATGACCGCTAAAACGGAAGGAG 1158
|||||
Db 355 AGACTCGAGTCTTTTCAAGCCAACTGCCACCATCATGACCGCTAAAACGGAAGGAG 296
|||||
QY 1159 ACTTCGGATAAAACAGCAGGAGCTGCGAACAAAGAAACAAAGGCTGTTGGAAGAAAG 1218
|||||
Db 295 ACTTCGGATAAAACAGCAGGAGCTGCGAACAAAGAAACAAAGGCTGTTGGAAGAAAG 236
|||||
QY 1219 AAATAATCTTGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278
|||||

```

```

Db 235 AAATAATCTTGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 176
|||||
QY 1279 TTGCTTTAGATATTATTAACTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTCTGCTCATG 1338
|||||
Db 175 TTGCTTTAGATATTATTAACTCCCTGTTTAACTCAGAGCTTTGGTAAAGTTTCTGCTCATG 116
|||||
QY 1339 TTCAAGCTGGGTAAGTTAGTTGTTGTTGAAGAGATTTGGTGTACCAAGTAAACAACTT 1398
|||||
Db 115 TTCAAGCTGGGTAAGTTAGTTGTTGTTGAAGAGATTTGGTGTACCAAGT-ACAAGAACTT 57
|||||
QY 1399 ATCGCTGTTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1431
|||||
Db 56 ATCGCTGTTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24
|||||

RESULT 15
BM501417 455 bp mRNA linear EST 14-FEB-2002
LOCUS PAC000000000714 Pioneer AF-1 array Zea mays cDNA, mRNA sequence.
DEFINITION BM501417
ACCESSION BM501417
VERSION BM501417.1 GI:18661475
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE clade; Panicoideae; Andropogoneae; Zea.
JOURNAL 1 (bases 1 to 455)
COMMENT Hunter, B.G., Beatty, M., Singletary, G., Hamaker, B., Larkins, B.A. and
Jung, R.
Maize opaque endosperm mutations create extensive changes in
patterns of gene expression
Unpublished (2002)
Contact: Jung R
Trait and Technology Development, Food and Feed Research
Pioneer Hi-Bred International, Inc.
7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
Tel: 515 270 5934
Fax: 515 254 2619
Email: rudolf.jung@pioneer.com.
Location/Qualifiers
1. 455
/organism="Zea mays"
/db_xref="taxon:4577"
/clone_lib="Pioneer AF-1 array"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI"
BASE COUNT 148 a 81 c 114 g 111 t 1 others
ORIGIN
Query Match 29.0%; Score 423.6; DB 13; Length 455;
Best Local Similarity 98.5%; Pred. No. 5.9e-68;
Matches 448; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 291 CAACGGACATAGATTACTTGAAGCGGGAATCAAGCAGTTTATGTTTGGATGGCAA 350
|||||
Db 2 CAACGGACATAGATTACTTGAAGCGGGAATCAAGCAGTTTATGTTTGGATGGCAA 61
|||||
QY 351 GCCTCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAAC 410
|||||
Db 62 GCCTCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAAC 121
|||||
QY 411 CAAAGATCTGACTGAGGAGTAGAGGTAGGAGATAAGATCGGATTGAAAATTCAGCAA 470
|||||
Db 122 CAAAGATCTGACTGAGGAGTAGAGGTAGGAGATAAGATCGGATTGAAAATTCAGCAA 181
|||||
QY 471 GAGGACTGTAAGCTCAAGGCAACACAGAGATTGTAAACGGCTATTAAAGACTTAT 530
|||||
Db 182 GAGGACTGTAAGCTCAAGGCAACACAGAGATTGTAAACGGCTATTAAAGACTTAT 241
|||||
QY 531 GGGGTTCTCTGTTAGAGGCACTTCTTGAGCAGAGAGAGAGATGTGAGCCCTTTGAT 590
|||||
Db 242 GGGGTTCTCTGTTAGAGGCACTTCTTGAGCAGAGAGAGAGATGTGAGCCCTTTGAT 301
|||||

```

```

Qy 591 AAACGATAAGGTGTTGCTGTTGCTTCAGAAAGATATGGAATCCCTTACTTTTGGGGCTCC 650
Db 302 AAACGATAAGGTGTTGCTGTTGCTTCAGAAAGATATGGAATCCCTTACTTTTGGGGCTCC 361
Qy 651 ACGGTTCCCTTCGTCATTAAAT-GGATCCAAAGTTCCAAAGAAAATACCTGTGATGGAATTTG 709
Db 362 ACGGTTCCCTTCGTCATTAAATGGGATCCAAAGTTCCAAAGAAAATACCTGTGATGGAATTTG 421
Qy 710 ATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCAC 744
Db 422 ATGTTGCCAAGG-TTGGAGGAGCCTTAACTCANC 455

```

Search completed: May 8, 2003, 21:34:20
 Job time : 1973 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 19:58:42 ; Search time 62 Seconds
(without alignments)
814,548 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGIKGLTKLADNAPKAMKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1939	100.0	379	21	Maize Rad2/FEN-1 p
2	1939	100.0	379	21	Maize Rad2/FEN-1 p
3	1933	99.7	379	21	Maize Rad2/FEN-1 p
4	1933	99.7	379	21	Maize Rad2/FEN-1 p
5	1031	53.2	380	20	Human FEN-1 protei
6	1007.5	52.0	377	20	Mouse FEN-1 protei
7	998.5	51.5	365	22	Drosophila melanog
8	996	51.4	378	20	Human FEN-1 protei
9	938	48.4	382	20	Yeast FEN-1 protei
10	699.5	36.1	373	22	Novel human diagno

11	648.5	33.4	340	18	AAW24216	Pyrococcus furiosu
12	648.5	33.4	340	19	AAW79970	Pyrococcus furiosu
13	648.5	33.4	340	19	AAW59940	Amino acid sequenc
14	647	33.4	326	19	AAW59953	Amino acid sequenc
15	644	33.2	332	19	AAW59951	Amino acid sequenc
16	636	32.8	343	22	AAW96644	Putative P. abyssal
17	632.5	32.6	343	20	AAW03778	Pyrococcus heat re
18	609	31.4	325	19	AAW59950	Amino acid sequenc
19	600	30.9	340	19	AAW59949	Amino acid sequenc
20	562	29.0	340	19	AAW59952	Amino acid sequenc
21	560.5	28.9	326	18	AAW24215	Methanococcus jann
22	560.5	28.9	326	19	AAW59939	Amino acid sequenc
23	553.5	28.5	326	19	AAW79969	Methanococcus jann
24	538.5	27.8	336	19	AAW79982	Archaeoglobus fulg
25	538.5	27.8	336	19	AAW59946	Amino acid sequenc
26	525	27.1	328	19	AAW59948	Amino acid sequenc
27	411	21.2	258	19	AAW59947	Amino acid sequenc
28	340.5	17.6	386	20	AAW92507	Yeast delta-RAD2 p
29	336	17.3	1516	21	AAW18195	Plasmodium faicipa
30	311	16.0	96	23	ABP33304	Human nuclease-lik
31	257	13.3	726	22	ABW65192	Drosophila melanog
32	243.5	12.6	1236	22	ABW65297	Drosophila melanog
33	243.5	12.6	1257	22	ABW67273	Drosophila melanog
34	215.5	11.1	871	23	AAW48927	Thermophilic DNA p
35	215.5	11.1	871	23	AAW48928	Thermophilic DNA p
36	210	10.8	732	22	ABW71917	Drosophila melanog
37	208	10.7	872	20	AAW31815	Thermoanaerobacter
38	205	10.6	872	18	AAW26605	Thermoanaerobacter
39	205	10.6	872	20	AAW31812	Thermoanaerobacter
40	205	10.6	872	20	AAW31816	T. thermohydrosulf
41	184.5	9.5	291	20	AAW78466	T. caldophilus DNA
42	181	9.3	832	13	AAW22602	Taq polymerase enc
43	174.5	9.0	832	13	AAW22603	Taq polymerase enc
44	174.5	9.0	832	13	AAW22604	Taq polymerase enc
45	173.5	8.9	834	17	AAW99542	Thermus flavus DNA

ALIGNMENTS

RESULT 1

AAW95307
ID AAW95307 standard; Protein; 379 AA.

XX AC AAW95307;

XX DT 12-SEP-2000 (first entry)

XX DE Maize Rad2/FEN-1 protein.

XX KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting.

XX OS Zea mays.

XX PN WO200036109-A1.

XX PD 22-JUN-2000.

XX PF 16-NOV-1999; 99WO-US27147.

XX PR 15-DEC-1998; 98US-0112332.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Mahajan PB;

XX DR WPI; 2000-452026/39.

XX DR N-PSDB; AAA27923.

XX PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -

XX Claim 11; Page 71-72; 85pp; English.
 XX The present sequence is that of maize Rad2/FEN-1, as deduced from
 CC a cDNA clone (see AAA27923) isolated from maize line B73 immature
 CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which
 CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
 CC can be expressed in transgenic plant cells using conventional
 CC methods. The protein is involved in the regulation of DNA repair
 CC and recombination in plant systems and therefore may be used for
 CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 XX Sequence 379 AA;

Query Match 100.0%; Score 1939; DB 21; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1.4e-170;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRTAVDASMSIYQFLIYVGRGTMETLTNEAGE 60
 Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRTAVDASMSIYQFLIYVGRGTMETLTNEAGE 60

Qy 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
 Db 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120

Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLMGVPVVEAPSEAEACALCINDKVFVASED 180
 Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLMGVPVVEAPSEAEACALCINDKVFVASED 180

Qy 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLEBELTMDQFDLCLCGCDYCDISIK 240
 Db 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLEBELTMDQFDLCLCGCDYCDISIK 240

Qy 241 GIGGTALKLIRHQHSIESILENLNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 Db 241 GIGGTALKLIRHQHSIESILENLNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300

Qy 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSQGRLESFFKPTATTAPLKRKETS 360
 Db 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSQGRLESFFKPTATTAPLKRKETS 360

Qy 361 DKTSKAAANKTKAGGKKK 379
 Db 361 DKTSKAAANKTKAGGKKK 379

RESULT 2
 AAY95309
 ID AAY95309 standard; Protein; 379 AA.

XX AAY95309;
 AC AAY95309;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Maize Rad2/FEN-1 protein.
 XX
 KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
 KW endonuclease; exonuclease; DNA repair; gene targeting.
 XX
 OS Zea mays.

XX WO200036109-A1.

XX

PD 22-JUN-2000.
 XX
 PF 16-NOV-1999; 99WO-US27147.
 XX
 PR 15-DEC-1998; 98US-0112332.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Mahajan PB;
 XX
 DR WPI; 2000-452026/39.
 DR N-PSDB; AAA27925.
 XX
 PT Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
 PT recombination and repair in transgenic plants, e.g. for gene targeting
 PT and the production of male sterile plants -
 XX
 PS Example 1; Page 76-77; 85pp; English.
 XX

The present sequence is that of maize Rad2/FEN-1, as deduced from
 CC a cDNA clone (see AAA27925) derived from maize line W23 tassel
 CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which
 CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
 CC can be expressed in transgenic plant cells using conventional
 CC methods. The protein is involved in the regulation of DNA repair
 CC and recombination in plant systems and therefore may be used for
 CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 XX

XX Sequence 379 AA;

Query Match 100.0%; Score 1939; DB 21; Length 379;
 Best Local Similarity 100.0%; Pred. No. 1.4e-170;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRTAVDASMSIYQFLIYVGRGTMETLTNEAGE 60
 Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRTAVDASMSIYQFLIYVGRGTMETLTNEAGE 60

Qy 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
 Db 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120

Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLMGVPVVEAPSEAEACALCINDKVFVASED 180
 Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLMGVPVVEAPSEAEACALCINDKVFVASED 180

Qy 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLEBELTMDQFDLCLCGCDYCDISIK 240
 Db 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLEBELTMDQFDLCLCGCDYCDISIK 240

Qy 241 GIGGTALKLIRHQHSIESILENLNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 Db 241 GIGGTALKLIRHQHSIESILENLNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300

Qy 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSQGRLESFFKPTATTAPLKRKETS 360
 Db 301 PDEGLISFLVKDNGFNEDRVTKAIEKISAKNKSQGRLESFFKPTATTAPLKRKETS 360

Qy 361 DKTSKAAANKTKAGGKKK 379
 Db 361 DKTSKAAANKTKAGGKKK 379

RESULT 3

```
AAAY95308
ID AAY95308 standard; Protein; 379 AA.
XX
AC AAY95308;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 protein.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS Zea mays.
XX
PN WO200036109-A1.
XX
PD 22-JUN-2000.
XX
PF 16-NOV-1999; 99WO-US27147.
XX
PR 15-DEC-1998; 98US-0112332.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Mahajan PB;
XX
DR WPI; 2000-452026/39.
DR N-PSDB; AAA27924.
XX
Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
recombination and repair in transgenic plants, e.g. for gene targeting
and the production of male sterile plants -
XX
Example 1; Page 74; 85pp; English.
XX
The present sequence is that of maize RAD2/FEN-1, as deduced from
a cDNA clone (see AAA27924) derived from maize line B73 seedling
tissue RNA. Rad2/FEN-1 is a structure specific endonuclease which
under certain conditions also acts as an exonuclease. Rad2/FEN-1
can be expressed in transgenic plant cells using conventional
methods. The protein is involved in the regulation of DNA repair
and recombination in plant systems and therefore may be used for
improving gene targeting during further recombinant DNA protocols
involving plants. RAD2/FEN-1 endonucleolytic activity is essential
in DNA replication and nucleotide excision and repair reactions.
The exolytic activity is involved in double strand break repair and
end joining. The protein is also useful in strand exchange
reactions during homologous recombination. These functions may be
useful in gene targeting and in the production of male sterile
plants. The efficacy of gene targeting can be improved by the
overexpression of exogenous Rad2/FEN-1 while male sterile plants can
be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 379 AA;
Query Match 99.7%; Score 1933; DB 21; Length 379;
Best Local Similarity 99.7%; Pred. No. 5e-170;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGIGKUTKLLADNAPKAMKEQFESYFGRKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
|||||
DB 1 MGIGKGLTKLLADNAPKAMKEQFESYFGRKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
|||||
QY 61 VTSHLOGMNRTTLLFAGIKPYYVDGKPPDMKKQBLAKRYSKRDDATKDLTFAVEVGD 120
|||||
DB 61 VTSHLOGMNRTTLLFAGIKPYYVDGKPPDMKKQBLAKRYSKRDDATKDLTFAVEVGD 120
|||||
QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPPVVEAPSEAEACALCINDKVFVASED 180
|||||
DB 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPPVVEAPSEAEACALCINDKVFVASED 180
|||||
QY 181 MSLTFTGAPRFLRHLMDPSSKKIPVMEFDVAKVLELELTMDQFIDLCLCGDCYCDISIK 240
|||||
```


PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
PS
XX Claim 1; Fig 2A; 58pp; English.
XX This sequence represents a mouse FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 377 AA;
Query Match 52.0%; Score 1007.5; DB 20; Length 377;
Best Local Similarity 53.4%; Pred. No. 1.7e-84;
Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;
QY 1 MGIGLTKLLADNAPKAMKQKESYFGRKTAVDASMSIYQFLIVVGRGTMETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKQKESYFGRKTAVDASMSIYQFLIVVGRGTMETLTNEAGE 59
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDKMKQELAKRYSKRDDATKDLTEAVEVD 120
Db 60 TTS-LMGMFYRTIR-MENGIKPVYVFDGKPPDKMKQELAKRYSKRDDATKDLTEAVEVD 116
QY 121 KDAIEKLSKRYVYTRQHNEDCKRLLRLMGVPVVEAPSEAECAALCINDKVPFAVASED 180
Db 117 MEEVEKTKRLVYTKQHNEDCKRLLRLMGVPVVEAPSEAECAALCINDKVPFAVASED 176
QY 181 MDSITFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCDYCSIK 240
Db 177 MDCLTFGSPVLMRLHTASEAKKLPQIEFHLSRVQLQELGLNQEVDFCLILGSDYCESIR 236
QY 241 GIGGQALKLIRHSGSIESILENLNKDRIQIPEDMPYQEARLKEPNVT-LDIPELKWT 299
Db 237 GIGAKRAVDLIQKHKSIEEIVRRIDPSKYPVFNWHLKEAQLLEPEVDVDPESVELKWS 296
QY 300 APDEGLISFLVKGNGFNEEDRVTKAIEKISAKNKSQGRLESFFKPTATTSAPLKRKET 359
Db 297 EPNEELVKFCGKQFSEERISGVRLSKRSRQSTQGRLLDFKFTGSLSS-AKRKEP 355
QY 360 SDKTSKAAANKTKTAGKKK 379
Db 356 EPKGS---AKKKARTGGAGK 372
RESULT 7
AB63960
ID ABB63960 standard; Protein; 385 AA.
XX
AC ABB63960;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 18672.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX OS WO200171042-A2.
XX
XX PN 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX

PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL08063.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 18672; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA
CC sequences (AB101840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 385 AA;
Query Match 51.5%; Score 998.5; DB 22; Length 385;
Best Local Similarity 52.9%; Pred. No. 1.2e-83;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;
QY 1 MGIGLTKLLADNAPKAMKQKESYFGRKTAVDASMSIYQFLIVVGRGTMETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKQKESYFGRKTAVDASMSIYQFLIVVGRGTMETLTNEAGE 59
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDKMKQELAKRYSKRDDATKDLTEAVEVD 120
Db 60 PTSHLMGMFYRTIRLLDNGIKPVYVFDGKPPDKMKQELAKRYSKRDDATKDLTEAVEVD 119
QY 121 KDAIEKLSKRYVYTRQHNEDCKRLLRLMGVPVVEAPSEAECAALCINDKVPFAVASED 180
Db 120 DAGIEKFNRRIVRVYTKQHNEDCKRLLRLMGVPVVEAPSEAECAALCINDKVPFAVASED 179
QY 181 MDSITFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCDYCSIK 240
Db 180 MDALTFGSTKRLRYLTSEARKMPVKESYDKLLEGLAINNREFIDLCILGCDYCESIK 239
QY 241 GIGGQALKLIRHSGSIESILENLNKDRIQIPEDMPYQEARLKEPNVT-LDIPELKWT 299
Db 240 GIGPKRAIELINTYRDIETILDNDSSKYTVPENNNYKVAELFTEPEVADADSIDLKWV 299
QY 300 APDEGLISFLVKGNGFNEEDRVTKAIEKISAKNKSQGRLESFFK---PTATTSAPL 354
Db 300 EPDEGLVFLCGDRQFNEERVRNGAKMLKMSKQATQVRLDSFFKTLFSTPNATNAA-- 357
QY 355 KRKETSDDKTSKAAANKTKTAGG 376
Db 358 KRK--AEAKKSANNKKAATSG 377
RESULT 8
AAW92508
ID AAW92508 standard; Protein; 378 AA.
XX
XX AAW92508;
XX
XX 23-APR-1999 (first entry)
XX
XX Human FEN-1 protein.
XX
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
XX

KW neoplasma; antineoplastic agent; cleavage.
XX Homo sapiens.
OS
XX
PN US5874283-A.
XX
XX 23-FEB-1999.
PD
XX
XX 30-MAY-1995; 95US-0455968.
PF
XX
XX 30-MAY-1995; 95US-0455968.
PR
XX
XX (HARR/) HARRINGTON J J.
PA (HSIEH/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
DR N-PSDB; AAX02111.
DR
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
PT
XX
PS Disclosure; Fig 5A-B; 58pp; English.
XX
XX This sequence represents a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 378 AA;
Query Match 51.4%; Score 996; DB 20; Length 378;
Best Local Similarity 52.4%; Pred. No. 2e-83;
Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;
QY 1 MGIGLTKLLADNAPKAMKEQKFEYFGKRIADVADSMSTIYQFLIVVGRGTMTLTNEAGE 60
Db 1 MEIHGLALIDVAPSAIRENDIKSYFGKKVAIDASMSTIYQFLIAY-RQGGDVLOQNEEGE 59
QY 61 VTSLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 60 TTS-LMGMFYRTMR-MENGIRPVYVFDGKPPOLKSGELAKRSERRAEAKLOOQAQAGM 117
QY 121 KDAIEKLSKRVKVTROHNEDECKRLRLMGVVPVPAEAEAEACALCINDKVFVASED 180
Db 118 EEEVEKFKRLVKVTKQBNDECKHLLMGIPYLDPAEAEAEACALAKAGKGVYAAATED 177
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFDLCLCGCDYCDSTK 240
Db 178 MDCLTFGSPVLMRLHMTASEAKLPQIEFHLKSLVGLGELNDEQFVLDCLILGSDYCESIR 237
QY 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDWPYQEARLKEPNVT-LDIPELKWT 299
Db 238 GIGAKRAVDLIQHKHSIEEIVRLDPSKYPPVFNWHLKEAQQLFLEPEVDPESVELAWS 297
QY 300 APDEGLLSFLVKONGFNEDRVTKAIEKIKSAKNKSOGRLSEFPKPTATTSAPLKRKET 359
Db 298 EPNEEELVKFCGEGKQFFFEIRSGVKRLSKSROGSTQGRLLDFFKVTGSLSS-AKRKEP 356
QY 360 SDKTSKAAANKTKAGGKKK 379
Db 357 E---PKGPAKKAKATGGAGK 373
RESULT 9

AAW92506
ID AAW92506 standard; Protein; 382 AA.
XX
AC AAW92506;
XX
DT 23-APR-1999 (first entry)
XX
DE Yeast FEN-1 protein.
XX
KW FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage.
XX
OS Saccharomyces cerevisiae.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
XX 30-MAY-1995; 95US-0455968.
PF
XX
XX 30-MAY-1995; 95US-0455968.
PR
XX
XX (HARR/) HARRINGTON J J.
PA (HSIEH/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
DR N-PSDB; AAX02109.
DR
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
PT
XX
PS Disclosure; Fig 3A; 58pp; English.
XX
XX This sequence represents a yeast FEN-1 (flap endonuclease) protein. This
CC protein is used in a method to isolate novel human FEN-1 proteins for
CC detecting a pathological condition in a patient, for diagnostic purposes,
CC for screening for antineoplastic agents and carcinogens, for diagnostic
CC staging of neoplasia, for producing recombinant flap endonuclease for use
CC as research or diagnostic reagents, for producing antibodies reactive
CC with the novel polypeptides, for producing transgenic nonhuman animals
CC expressing the novel polypeptides encoded by a transgene. The invention
CC also provides novel molecular cloning techniques and reagents involving
CC cleavage of a flap or nick with a flap endonuclease.
XX
SQ Sequence 382 AA;
Query Match 48.4%; Score 938; DB 20; Length 382;
Best Local Similarity 49.5%; Pred. No. 4.6e-78;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;
QY 1 MGIGLTKLLADNAPKAMKEQKFEYFGKRIADVADSMSTIYQFLIVVGRGTMTLTNEAGE 60
Db 1 MGIGLNAIIIEHVPSAIRKSDIKSFYFGKKVAIDASMSTIYQFLIARQDDGQLTNEAGE 60
QY 61 VTSLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 TTS-LMGMFYRTLRMDNGIRPVYVFDGKPPDLKSHLTKRSRRVETEKLAETTELE 120
QY 121 KDAIEKLSKRVKVTROHNEDECKRLRLMGVVPVPAEAEAEACALCINDKVFVASED 180
Db 121 K---MKQERRLVKVSKHENEEAQKLLGLMGIPYIIAPTAEAAQCAELAKGKGVYAAASD 177
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFDLCLCGCDYCDSTK 240
Db 178 MDCLCYRTPFLRLHMTASEAKLPQIEFHLKSLVGLGELNDEQFVLDCLILGSDYCESIR 237
QY 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDWPYQEARLKEPNVT-LDIPET 295
Db 238 GVGPPVTKLTKTHGSIETKIVEFTESGNSNNTKWKIPEDWPYKQARMFLDPEV-IDGNE 296

QY 296 --LKWTPADEGLISFLYKDNFGNEDRVTKATEIKSAKNSOGRLSEFFKPTATT--- 350
 DB 297 INLWSPPEKELEIYLCDDKFSERVKSGISRLKGLKSGIQGRLDGFFQVVPKTEQ 356
 QY 351 -SAPLKRKETSDKTSAKAAKTK 373
 DB 357 LAAAKRAQENKLNK-NKNKVK 379

RESULT 10
 ABG19545
 ID ABG19545 standard; Protein; 373 AA.
 XX
 AC ABG19545;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19536.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS83732.
 XX
 PT New isolated polypeptide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 49904; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 373 AA;

Query Match 36.1%; Score 699.5; DB 22; Length 373;
 Best Local Similarity 41.1%; Pred. No. 4.9e-56;

Matches 155; Conservative 68; Mismatches 105; Indels 49; Gaps 7;

QY 1 MGIGITKLADNAPKAMKEQKESYFGPKIAVDASMSYQFLIVVGRGTMETLTNEAGE 60
 DB 36 MGIOGLAKLIADVAPSAIRENDIKSTYGRKVAIDASMSYHFLIAV-ROGGDVLQNEEGE 94
 QY 61 VTSHLQGMENRTIRLEAGIKPVPYVDFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
 DB 95 TTSHLMGMFYRTIRMKENGKIPY-----KSGELAKCSLRAEAEKQLQQAQAAGA 145
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPVVEPAEAECAALCINDKVFAYASED 180
 DB 146 EQVVEKFTKWLKVTQKHNCKHLLSLMGIPYDAIEAEASCAALVAKGVYPVATED 205
 QY 181 MDSLTFGAPRFLHLMDPSSKKIPVMEFDVAKVLELELTMDQFIDLCLICGGDYCDISK 240
 DB 206 MDCFTFGSPVLMOHLTASS-----CYCKSIW 231
 QY 241 GIGGOTALKLIHQHGSIESILENLNDRYOIPEDWPYQARRLFKEPNVT-LDIPELKWT 299
 DB 232 SIGPKRAVDLIQKHKSIEIIVQRLDPNKVPLPENRLHKEAYQLFLEPVLGPESVELKWS 291
 QY 300 APDEGLISFLYKDNFGNEDRVTKATEIKSAKNSOGRLSEFFKPTATTTSAPLKRKE- 358
 DB 292 EPNEEELVKFVCGEQKOFSEERIHSGVKRLTEKG-GGEPRGPDWTQP--GRASFTKRPDH 348
 QY 359 TSDKTSKAAANKTKAG 375
 DB 349 QEDQHTRSKSGSHXG 365

RESULT 11
 AAW24216
 ID AAW24216 standard; Protein; 340 AA.
 XX
 AC AAW24216;
 XX
 DT 14-APR-1998 (first entry)
 XX
 DE Pyrococcus furiosus FEN-1 endonuclease.
 XX
 KW Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
 KW invader directed cleavage; FEN-1; endonuclease.
 XX
 OS Pyrococcus furiosus.
 XX
 PN WO9727214-A1.
 XX
 PD 31-JUL-1997.
 XX
 PF 22-JAN-1997; 97WO-US01072.
 XX
 PR 02-DEC-1996; 96US-0759038.
 PR 24-JAN-1996; 96US-059491.
 PR 12-JUL-1996; 96US-0882853.
 PR 29-NOV-1996; 96US-0756386.
 PR 02-DEC-1996; 96US-0758314.
 XX
 (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;
 PI Olive DM, Prudent JR;
 XX
 DR WPI; 1997-393613/36.
 DR N-PSDB; AAT76685.
 XX
 PT Thermostable structure-specific nuclease(s) - used for detection and
 PT characterisation of nucleic acid sequences and variations in nucleic
 PT acid sequences
 XX
 PS Example 28; Page 283-285; 457pp; English.
 XX
 CC This sequence comprises Pyrococcus furiosus (Pfu) FEN-1

CC endonuclease. Large-scale production of the enzyme was performed
 CC using *E. coli* host cells transformed with a vector carrying the Pfu
 CC FEN-1 coding sequence (see AAT76685). Pfu FEN-1 is a thermostable
 CC enzyme. It can be used in novel methods for the detection and
 CC characterisation of nucleic acid sequences and variations in
 CC nucleic acid sequences.

XX Sequence 340 AA;

Query Match 33.4%; Score 648.5; DB 18; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.2e-51;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KQKESYEGRIKIVDASMSIQFLIVGRTGMEITNAGEVTSHTLQGMFNRITLLEA 78
 DB 12 KEIELENLYGKKIADALNAIYQFLSTIRQKDGTPLMDSKGRITSHLSGLFYRTINLMEA 71
 QY 79 GKPYVYFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKALEKLSKRTVKVTRQH 138
 DB 72 GKPYVYFDGEPPEPKKELEKREAREAEKWEAEKEIEEARKYAQATRVNEML 131
 QY 139 NEDCKRLRLMGVPVPEAPSEAEACALCINDKVFVASEDMDSITFGAPRFLRLMDP 198
 DB 132 IEDAKKLLLMGIPITVQAPSEGAQAAYMAAKGSVYASQDYDLSLLFGAPRLVRNLIT 191
 QY 199 SSKKIP-----VMEFDVAKVLELELTMDQFIDILCGCDY-CDSIKGIGG 244
 DB 192 GKRKLPGKNVYVEIKPELIILE-----EVLKELKLTREKLIELAILVGTDPNGGKIGL 247
 QY 245 QTALKLRHQHSIESILENLNKD---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWAP 301
 DB 248 KKALEIVRH-----SKDPLAKFKQSDVDLYAIKEFFLNPPVT-DNYNLVWRDP 295
 QY 302 DEGLISFLVNDGNEDRVTKAEIKSAKNKSSQGRLESFFK 345
 DB 296 DEEGILKFLCDEHDFSEERVKNGLERLKAIAKSKQSTLESWFK 339

RESULT 12

AAW79970
 XX AAW79970 standard; Protein; 340 AA.
 AC AAW79970;
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Pyrococcus furiosus FEN-1 endonuclease.
 XX Nucleic acid detection; multiple sequential invasive cleavage;
 KW FEN-1; endonuclease; nuclease.
 XX Pyrococcus furiosus.
 OS
 XX WO9842873-A1.
 XX
 XX 01-OCT-1998.
 XX
 XX 24-MAR-1998; 98WO-US05809.
 XX
 XX 24-MAR-1997; 97US-0823516.
 XX
 XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 XX BROW MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
 PI Mast AL, Vavra SH;
 XX WPI; 1998-557036/47.
 DR N-PSDB; AAV65840.
 XX

Detecting target nucleic acid by sequence-specific cleavage of
 PT complex with two specific oligonucleotides - used to detect
 PT cytomegalovirus DNA
 XX

PS Example 28b; Page 316-317; 524pp; English.

XX This is the amino acid sequence of FEN-1 endonuclease of
 CC Pyrococcus furiosus (Pfu). FEN-1 DNA (see AAV65840) has been
 CC ligated into vector for expression in *E. coli* cells. The invention
 CC relates to means for the detection and characterisation of nucleic
 CC acid sequences, and variations in nucleic acid sequences. It also
 CC relates to methods for forming a nucleic acid cleavage structure on
 CC a target sequence and cleaving this structure in a site-specific
 CC manner, preferably using a thermostable structure-specific nuclease
 CC such as FEN-1. Cleavage of the cleavage structure by the
 CC nuclease indicates the presence of specific nucleic acid sequences
 CC or specific variants. The invention further relates to methods for
 CC the separation of nucleic acid molecules based on charge, methods
 CC for the detection of non-target cleavage products via the formation
 CC of a complete and activated protein binding region, and methods for
 CC the detection of nucleic acid from various viruses (e.g. human
 CC cytomegalovirus) in a sample. The method amplifies the detection
 CC molecule rather than the target itself, is less subject to
 CC contamination than exponential amplification processes, and allows
 CC many targets to be analysed in a single reaction.

XX Sequence 340 AA;

Query Match 33.4%; Score 648.5; DB 19; Length 340;
 Best Local Similarity 41.3%; Pred. No. 2.2e-51;
 Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KQKESYEGRIKIVDASMSIQFLIVGRTGMEITNAGEVTSHTLQGMFNRITLLEA 78
 DB 12 KEIELENLYGKKIADALNAIYQFLSTIRQKDGTPLMDSKGRITSHLSGLFYRTINLMEA 71
 QY 79 GKPYVYFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKALEKLSKRTVKVTRQH 138
 DB 72 GKPYVYFDGEPPEPKKELEKREAREAEKWEAEKEIEEARKYAQATRVNEML 131
 QY 139 NEDCKRLRLMGVPVPEAPSEAEACALCINDKVFVASEDMDSITFGAPRFLRLMDP 198
 DB 132 IEDAKKLLLMGIPITVQAPSEGAQAAYMAAKGSVYASQDYDLSLLFGAPRLVRNLIT 191
 QY 199 SSKKIP-----VMEFDVAKVLELELTMDQFIDILCGCDY-CDSIKGIGG 244
 DB 192 GKRKLPGKNVYVEIKPELIILE-----EVLKELKLTREKLIELAILVGTDPNGGKIGL 247
 QY 245 QTALKLRHQHSIESILENLNKD---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWAP 301
 DB 248 KKALEIVRH-----SKDPLAKFKQSDVDLYAIKEFFLNPPVT-DNYNLVWRDP 295
 QY 302 DEGLISFLVNDGNEDRVTKAEIKSAKNKSSQGRLESFFK 345
 DB 296 DEEGILKFLCDEHDFSEERVKNGLERLKAIAKSKQSTLESWFK 339

RESULT 13

AAW59940
 ID AAW59940 standard; Protein; 340 AA.
 XX AAW59940;
 XX
 XX 21-DEC-1998 (first entry)
 XX
 XX Amino acid sequence of the Pfu FEN-1 endonuclease.
 XX Mja FEN-1 endonuclease; Taq gene; structure-specific nuclease;
 KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;
 KW hepatitis C virus; HCV; thermostable.
 XX Pyrococcus furiosus.
 OS

WO9823774-A1.
 XX
 XX 04-JUN-1998.
 XX

DE Amino acid sequence of the structure specific nuclease 3.
 KW Taq mutant gene; thermostable; nuclease; mutant; DNA polymerase;
 KW bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
 XX
 XX Synthetic.
 OS
 XX
 XX WO9823774-A1.
 PN
 XX
 XX 04-JUN-1998.
 PD
 XX
 XX 26-NOV-1997; 97WO-US21783.
 PF
 XX
 XX 02-DEC-1996; 96US-0758314.
 PR
 XX 29-NOV-1996; 96US-0757653.
 PR
 XX
 XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
 PA
 XX
 XX Kaiser MW, Lyamichev VI, Lyamicheva N;
 PI
 XX
 XX WPI; 1998-322748/28.
 DR
 XX
 XX Thermostable structure-specific nuclease(s) derived from mutant DNA
 PT polymerase(s) - useful for detecting mutant allele(s) or strains of
 PT microorganisms
 PT
 XX
 XX Claim 1; Pages 329-330; 472pp; English.
 PS
 XX
 XX This is the amino acid sequence of a structure specific nuclease
 CC used in the method of the invention. In this process thermostable
 CC structure-specific nucleases are derived from mutant DNA polymerases,
 CC which can be used for detecting mutant alleles or strains of
 CC microorganisms. The structure-specific nucleases can be used in
 CC mixtures, compositions and kits to treat nucleic acid, e.g. for
 CC detection of wild type and mutant alleles of genes, for detection
 CC and/or identification of strains of microorganisms such as bacteria,
 CC fungi, protozoa, especially for detection of RNA viruses such as the
 CC hepatitis C virus (HCV).
 CC
 XX
 XX Sequence 332 AA;
 SQ

Query Match 33.2%; Score 644; DB 19; Length 332;
 Best Local Similarity 43.2%; Pred. No. 5.6e-51;
 Matches 142; Conservative 68; Mismatches 103; Indels 16; Gaps 7;

QY 19 KEOKFESYFGKRIADVDSMSIYQFLIVVGRGTGMEITLTNEAGEVTSHLQGMFNTIRLLEA 78
 Db 12 KELELENLGKKAIDALNAYIQLSTIRKQDGTPLMDSKGRITSHLSGLFYRTINLMEA 71

QY 79 GIKPVYVDFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKDATEKLSKRTVKVTRQH 138
 Db 72 GIKPVYVDFGEPPEFKKLEKREAREAEKREALEKEGEIEEAAKAKRVSYLTPRM 131

QY 139 NEDCKRLRLMGVPPVVEAPSEAEACALCINDKVFVASEDMDSLTFGAPREFRLHMDP 198
 Db 132 VENCKYLLSLMGIPYVEAPSEGAQAQYNAKKGDMVAVSQDYDALLYGAPRVVRNL--T 189

QY 199 SSKIPVMEFDVAKVLEELTMDQFDLCLCGCDY-CDSIKGIGGQTKLIRHGS 257
 Db 190 TTKEMPEL-IELNEVLDELRLISLDDLDIAFMGTDFNPGGVKYGIGKRAYELVRS-GVA 247

QY 258 ESILENLNDRYQIPEDWPYQEARLFKEPNVTLDIPELKWTAPDEGLISFLVKDNGFN 317
 Db 248 KDVLLK-----EVEYYDEIKRIFKEPKVT-DNYSLSKLDPKREGIIFKFLVDENDFN 297

QY 318 EDRTVTAIEKIKS-AKNKSSQGRLESFFK 345
 Db 298 YDRVKKHVDKLYNLANKTKQKTLDAWFK 326

Search completed: May 8, 2003, 21:40:45
 Job time : 64 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:39:38 ; Search time 34 Seconds
(without alignments)
327.979 Million cell updates/sec

Title: US-09-805-311-2
Perfect score: 1939
Sequence: 1 MGKGLTKLLADNAPKAMKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfilesi.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1939	100.0	379	4	US-09-426-557-2
2	1939	100.0	379	4	US-09-426-557-6
3	1933	99.7	379	4	US-09-426-557-4
4	1933	99.7	379	4	US-09-426-557-8
5	1032	53.2	380	2	US-08-823-516-137
6	1032	53.2	380	2	US-09-426-557-10
7	1031	53.2	380	2	US-08-455-968E-1
8	1014	52.3	378	2	US-08-823-516-138
9	1007.5	52.0	377	2	US-08-455-968E-3
10	996	51.4	378	2	US-08-455-968E-10
11	938	48.4	382	2	US-08-455-968E-5
12	938	48.4	382	2	US-08-823-516-139
13	648.5	33.4	340	2	US-08-757-653-176
14	648.5	33.4	340	2	US-08-823-516-79
15	648.5	33.4	340	2	US-08-823-516-136
16	648.5	33.4	340	3	US-08-759-038-115
17	648.5	33.4	340	3	US-08-314-115
18	632.5	32.6	343	4	US-09-146-319-2
19	632.5	32.6	343	4	US-09-175-973-2
20	560.5	28.9	326	2	US-08-757-653-172
21	560.5	28.9	326	2	US-08-823-516-75
22	560.5	28.9	326	2	US-08-823-516-135
23	560.5	28.9	326	3	US-08-759-038-111
24	560.5	28.9	326	3	US-08-314-111
25	364	18.8	387	2	US-08-823-516-140
26	352.5	18.2	488	2	US-08-823-516-141
27	340.5	17.6	386	2	US-08-455-968E-7

28	340	17.5	543	2	US-08-823-516-143	Sequence 143, Appl
29	334.5	17.3	550	2	US-08-823-516-142	Sequence 142, Appl
30	313.5	16.2	527	2	US-08-823-516-144	Sequence 144, Appl
31	205	10.6	872	1	US-08-766-014-2	Sequence 2, Appli
32	184.5	9.5	291	4	US-09-105-697-6	Sequence 6, Appli
33	173.5	8.9	834	5	PCT-US95-14418-2	Sequence 2, Appli
34	173.5	8.9	834	5	PCT-US95-15327-2	Sequence 2, Appli
35	173	8.9	548	2	US-08-484-956-86	Sequence 86, Appl
36	173	8.9	548	2	US-08-757-653-86	Sequence 86, Appl
37	173	8.9	548	2	US-08-520-946-86	Sequence 86, Appl
38	173	8.9	695	2	US-08-484-956-87	Sequence 87, Appl
39	173	8.9	695	2	US-08-757-653-87	Sequence 87, Appl
40	173	8.9	810	4	US-08-520-946-87	Sequence 2, Appli
41	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appli
42	173	8.9	810	4	US-09-777-537-2	Sequence 2, Appli
43	173	8.9	810	4	US-09-777-538-2	Sequence 2, Appli
44	173	8.9	832	1	US-07-977-434-2	Sequence 2, Appli
45	173	8.9	832	1	US-08-156-020-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-426-557-2
; Sequence 2, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-2

Query Match	100.0%	Score 1939;	DB 4;	Length 379;
Best Local Similarity	100.0%	Pred No. 1.2e-183;		
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGKGLTKLLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVGRTGTMETLTNEAGE	60	
Db	1	MGKGLTKLLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVGRTGTMETLTNEAGE	60	
QY	61	VTSHLOGMNRTRIRLEAGIKPVYVFDGKPPDMKKQBLAKRYSKRDATAKDLTAEVGD	120	
Db	61	VTSHLOGMNRTRIRLEAGIKPVYVFDGKPPDMKKQBLAKRYSKRDATAKDLTAEVGD	120	
QY	121	KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPPVVEAPSEAECAALCINDKVFVASED	180	
Db	121	KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPPVVEAPSEAECAALCINDKVFVASED	180	
QY	181	MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCYCDSIK	240	
Db	181	MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCYCDSIK	240	
QY	241	GIGGQTALKLIHQHGSIESILENLNDRYQIPEDWPQEARRLFKEPNVTLDPILKWT	300	
Db	241	GIGGQTALKLIHQHGSIESILENLNDRYQIPEDWPQEARRLFKEPNVTLDPILKWT	300	
QY	301	PDEGLISFLVDKNGFNEDRVTKAIEKISAKNKSOGRLSEFPKPTATTAPLKRKETS	360	
Db	301	PDEGLISFLVDKNGFNEDRVTKAIEKISAKNKSOGRLSEFPKPTATTAPLKRKETS	360	
QY	361	DKTSKAAANKTKAGGKKK	379	

Db 121 KDAEKLKSKTVKTRHNECDKRLRLMGVPPVVEAPSEAECAALCINDKVFVASED 180
QY 181 MDSLTGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELMDQFIDICILCGGDCYDSIK 240
Db 181 KDSLTGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELMDQFIDICILCGGDCYDSIK 240
QY 241 GIGGOTALKLIRHGSTESILENLDKRYOIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
Db 241 GIGGOTALKLIRHGSTESILENLDKRYOIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
QY 301 PDEEGLISFLVKGNGFNEEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
Db 301 PDEEGLISFLVKGNGFNEEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAAANKTKRAGGKKK 379
Db 361 DKTSKAAANKTKRAGGKKK 379

RESULT 5

US-08-823-516-137
; Sequence 137, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlin & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-823-516-137

Query Match 53.2%; Score 1032; DB 2; Length 380;
Best Local Similarity 53.8%; Pred. No. 7.1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
QY 1 MGIKGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTMTTNEAGE 60
Db 1 MGIOGLAKLIADVAPSAIRENDIKSYFGKRVKVAIDASMSIYQFLIIV-RGGGDLVQNEEGE 59
QY 61 VTSHLQGMFNRTIRLLEAGIKPVYFDGKPPDMKQELAKRYSKRDDATKDLTTEAVEVGD 120
Db 60 TTHLMGMFYRTIRMMENGIRPVYFDGKPPQLKSGELAKRSERRAEAKQLQQAAGA 119
QY 121 KDAEKLKSKTVKTRHNECDKRLRLMGVPPVVEAPSEAECAALCINDKVFVASED 180
Db 120 EQVEKFTKRLVKVTKOHNECKHLLSLMGIPYLDAPSEASCAALVKAGKVTAAATED 179
QY 181 MDSLTGAPRFLRLHMDPSSKKIPVMEFDVAKVLELELMDQFIDICILCGGDCYDSIK 240
Db 180 MDCITFGSPVLMRHLTASEAKKLPIDQEHLSRIILQELGLNQEQQFVDCILLAGSDYCESIR 239
QY 241 GIGGOTALKLIRHGSTESILENLDKRYOIPEDWPYQEARRLFKEPNVTLDIPELKW 298
Db 240 GIGPKRAVDLIQKHKSIEIVRRDPNKPYPENWLHKEAHQLEPEV-LDPESVELKW 298
QY 299 TAPDEEGLISFLVKGNGFNEEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKE 358
Db 299 SEPNEEELIKFMCGEKOFSEERISRGVRLSKSGTQGRDLDDFFKVTGSLSS-AKAKE 357
QY 359 TSDK--TSKAAANKTKRAGGKKK 379
Db 358 PEPKSGSTKKA--KTGAAGKFK 377

RESULT 6

US-09-426-557-10
; Sequence 10, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-426-557-10

Query Match 53.2%; Score 1032; DB 4; Length 380;
Best Local Similarity 53.8%; Pred. No. 7.1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
QY 1 MGIKGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTMTTNEAGE 60
Db 1 MGIOGLAKLIADVAPSAIRENDIKSYFGKRVKVAIDASMSIYQFLIIV-RGGGDLVQNEEGE 59
QY 61 VTSHLQGMFNRTIRLLEAGIKPVYFDGKPPDMKQELAKRYSKRDDATKDLTTEAVEVGD 120

Db 60 TTSHLMGFYRTIRMENCIKPVYVDFDCKPOLKSGELAKRSERRAEAEKQLQQAAGA 119
QY 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPPVVEAPSEAEAEALCINDKVFVASED 180
Db 120 EQEVEKFTKRLVKVTKQHNDCKHLLSLMGIPYLDAPSEAEASCAALVKAGKRYAAATED 179
QY 181 MDSLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFTDLCILCCGDCYCDSTK 240
Db 180 MDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILOELGQLNQEQFYDLCLILGSDYCESIR 239
QY 241 GIGGOTALKLIRHQHSIESILENKNDRYQIPEDWPYOEARLFKEPNVTLD--IPELKW 298
Db 240 GIGPRAVDLIQQHKHSIEIIVRLDPNKPYPENWLHLKEAHLQFLEPEV-LDPESVELKW 298
QY 299 TAPDEGLISFLVNDGNEEDRVTKAIEKISAKNKSQGRLESFFKPTATTSAPLKRKE 358
Db 299 SEPNEELIKFMCGEKQSEERIRSGVRLSKSRQSGTQGRDLDFFKVTGSLSS-AKRKE 357
QY 359 TSDK--TSKAAANKTKAGGKKK 379
Db 358 PEPKSTKKA---KTGAAGKFK 377

RESULT 7

US-08-455-968E-1

; Sequence 1, Application US/08455968E

; Patent No. 5874283

; GENERAL INFORMATION:

; APPLICANT: Harrington, John L.

; APPLICANT: Hsieh, Chih-Lin

; APPLICANT: Lieber, Michael

; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,968E

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.

; REGISTRATION NUMBER: 36,429

; REFERENCE/DOCKET NUMBER: 18985-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 380 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-455-968E-1

Query Match

Best Local Similarity 53.2%; Score 1031; DB 2; Length 380;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGLKLKLLADNAPKAMKQKFSYGRKIVADMSIYQFLIVGRTGNETITNAGE 60

Db 1 MGIGLKLKLLADNAPKAMKQKFSYGRKIVADMSIYQFLIVGRTGNETITNAGE 59

QY 61 VTSHLQGMFNRTIRLERAGIKPVYVDFDCKPOLKSGELAKRSERRAEAEKQLQQAAGA 120
Db 60 TTSHLMGFYRTIRMENCIKPVYVDFDCKPOLKSGELAKRSERRAEAEKQLQQAAGA 119
QY 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPPVVEAPSEAEAEALCINDKVFVASED 180
Db 120 EQEVEKFTKRLVKVTKQHNDCKHLLSLMGIPYLDAPSEAEASCAALVKAGKRYAAATED 179
QY 181 MDSLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFTDLCILCCGDCYCDSTK 240
Db 180 MDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILOELGQLNQEQFYDLCLILGSDYCESIR 239
QY 241 GIGGOTALKLIRHQHSIESILENKNDRYQIPEDWPYOEARLFKEPNVTLD--IPELKW 298
Db 240 GIGPRAVDLIQQHKHSIEIIVRLDPNKPYPENWLHLKEAHLQFLEPEV-LDPESVELKW 298
QY 299 TAPDEGLISFLVNDGNEEDRVTKAIEKISAKNKSQGRLESFFKPTATTSAPLKRKE 358
Db 299 SEPNEELIKFMCGEKQSEERIRSGVRLSKSRQSGTQGRDLDFFKVTGSLSS-AKRKE 357
QY 359 TSDK--TSKAAANKTKAGGKKK 379
Db 358 PEPKSTKKA---KTGAAGKFK 377

RESULT 8

US-08-823-516-138

; Sequence 138, Application US/08823516

; Patent No. 5994069

; GENERAL INFORMATION:

; APPLICANT: Hall, Jeff G.

; APPLICANT: Lyamichev, Victor I.

; APPLICANT: Mast, Andrea L.

; APPLICANT: Brow, Mary Ann D.

; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple

; TITLE OF INVENTION: Sequential Invasive Cleavages

; NUMBER OF SEQUENCES: 163

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States Of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/823,516

; FILING DATE: 24-MAR-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US97/01072

; FILING DATE: 21-JAN-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/759,038

; FILING DATE: 02-DEC-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/758,314

; FILING DATE: 02-DEC-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/756,386

; FILING DATE: 29-NOV-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/682,853

; FILING DATE: 12-JUL-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/599,491

; FILING DATE: 24-JAN-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Ingolia, Diane E.

; CLASSIFICATION: 435

; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-968E-10

Query Match 51.4%; Score 996; DB 2; Length 378;
Best Local Similarity 52.4%; Pred. No. 2.6e-90;
Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;
QY 1 MGIGIKTLADNAPKAMKEQKFESYFGKRIADVADSMSTYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MEIHLAKLIADVAPSAIRENDIKSYFGKRVADSMSTYQFLIIV-RGGDVQLQNEEGE 59
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVDFGKPPDMKKQELAKRYSKRDDATKLTAEVVG 120
Db 60 TTS-LMGFMFYRMR-MENGIKPVYVDFGKPPQKSGELAKRSERRAEKQLQQAQAGM 117
QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVAPSEAEACALCINDKVFVAVASD 180
Db 118 EEEVKFTKRLVKVTRQHNEDCKRLRLMGVVPVAPSEAEACALCINDKVFVAVASD 177
QY 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLELELTMDQFIDLCILCGCDYCDSTK 240
Db 178 MDCLTFGSPVLMRLHRTASEAKKLIQEPHLSVLQELGVLNQFVDFDLCLILGSDYCESIR 237
QY 241 GIGGOTALKLIRHQSIESILENLNKRQYQIPEDWPYQEARLFKEPNVT-LDIPELKWT 299
Db 238 GIGAKRAVDLIQKHSIEIVRRLDPSKYPVPENLHKEAQLFLEPVVDPSPESVELKWS 297
QY 300 APDEGLISFLVKDNGFNDRVTAKIEIKSAKNKSSQGRLESFFKPTATTAPLKRKET 359
Db 298 ENEBELVKFMGEKQFEEFRTSGVKRLSKRSQSTQGRLLDDFFKVTGSLSS-AKRREP 356
QY 360 SDKTSKAAANKTKKAGKKK 379
Db 357 E---PKGPAKKKAKTGGAGK 373

RESULT 11
US-08-455-968E-5
; Sequence 5, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-968E-5
Query Match 48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 1.4e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;
QY 1 MGIGIKTLADNAPKAMKEQKFESYFGKRIADVADSMSTYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MGIGLNLAIISHVPSAIRKSDIKSFFGKRVADSMSTYQFLIIVRQDGGQLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVDFGKPPDMKKQELAKRYSKRDDATKLTAEVVG 120
Db 61 TTS-HLMGMFYRTRMIDNGIKPCYVDFGKPPDKLSHETKRSRRVETKELAEATTELE 120
QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVAPSEAEACALCINDKVFVAVASD 180
Db 121 K---MKQERLLVKVSKHEEAEQKLLGLMGIPIIAPTAEAEQAQELAKKGVYAAASD 177
QY 181 MDSLTFGAPRFLRHLMDSPPSKKIPVMEFDVAKVLELELTMDQFIDLCILCGCDYCDSTK 240
Db 178 MDTLCYRPFLLRHITFSEAKKEPIHEIDTVELVGLDITIEQFVDLCIMLGCDCYCESIR 237
QY 241 GIGGOTALKLIRHQSIESILENLNKRQYQIPEDWPYQEARLFKEPNVTLDIPE 295
Db 238 GVGVPVTLKLIKTHGSIEKIVFIESGESNNTKWKIPEDWPYQARMFLDPEV-IDGNE 296
QY 296 --LKWTPADEGLISFLVKDNGFNDRVTAKIEIKSAKNKSSQGRLESFFKPTATT--- 350
Db 297 INKWSPPKRELIEYLCDDKKFSERVKSGISRLKGLKLSGIQGRLDGFFQVVPKTKEQ 356
QY 351 -SAPLKRKETS DKTSKAAANKTKK 373
Db 357 LAAAKRAQENKKLNK-NKNKYTK 379

RESULT 12
US-08-823-516-139
; Sequence 139, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-139

Query Match 48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 1.4e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;
Qy 1 MGKGLTKLLADNAPKAMEQKESFYGRKTIADVADSMISYQFLIVVGTGTMETLTNEAGE 60
Db 1 MGKGLNATISSEHPSPAIKKSIDKSFGRKVAIDASMSLYQFLTAVRQDGGQLTNEAGE 60
Qy 61 VTSLQGMENRTIRLLEAGIKPVYFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 TTSLMGMYRTLRMDNGIKPCYVFDGKPPDLKSHLTKSSRRVETKRLAEATTELE 120
Qy 121 KDAELKSKRTVATVTOHNDCKRLRLMGVPEAPSEAECAALCINDKVFVASED 180
Db 121 K--MKQERLVVSKHEHEAQLGLMGIPYITAPTEAQAQCAELAKGKGVYAAASED 177
Qy 181 MDSITFGAPRLHLMPPSSKKIPVMEFDVAKVLEELTMDQFTDLCILCGDCYDSIK 240
Db 178 MDTLCYCTPPLRLHRTFSKAKPEITHEIDTLVRLGLDTIEQFVDCILMGDCYCSIR 237
Qy 241 GIGQOTALKLIRHGSIETLENL-----NKDRYQIPEDWPYQEARRLFKEPNVTLDIPE 295
Db 238 GVGPTALKLKHGSIETLVERFIESGESNNTKKWIPEDWPYQKARMLFLDPEV-IDGNE 296
Qy 296 --LKWTPADEGLISFLVNDGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATT--- 350
Db 297 INLWSPKKEILYELCDDKFSERVKSGISRLKGLKSGIOGRULDGFQVVPKTKQ 356
Qy 351 -SAPLKRKETSDKTSAKAAKTK 373
Db 357 LAAAKRAQENKKLNK-NKNKVTK 379

RESULT 13
US-08-757-653-176
Sequence 176, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-176
Query Match 33.4%; Score 648.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 5.4e-56;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
Qy 19 KEQFESYFGKTIADVADSMISYQFLIVVGTGTMETLTNEAGEVTSHLQGMFNRITRLEA 78
Db 12 KEIELENYGKTIADNATYQFLSTIRQDKGTPLMDSKGRITSHLSGLFYRTINLMEA 71
Qy 79 GIKPVYFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKDAIEKLSKRTVKVTRQH 138
Db 72 GIKPVYFDGPEPEFKKLEKREAREAEKREALEKGEIEARVYQARATFVNEML 131
Qy 139 NEDCKRLRLMGVPEAPSEAECAALCINDKVFVASEDMSLTGAPRFLHLMDP 198
Db 132 IEDAKKLELLELGMIPVQAPSEGEQAAYMAAKSGSVYASQDYDLSLFGAPRLVRLTIT 191
Qy 199 SSRKIP-----VMEFDVAKVLELELTMDQFTDLCILCGDY-CDSIKGIGG 244
Db 192 GKRKLPGKNVVEETKPELITL-----EVLKELTKREKLIETLAILVGTDPNGIKIGL 247
Qy 245 OTALKLIRHGSIETLENLNDK---RYOIPEDWPYQEARRLFKEPNVTLDIPELKWTP 301
Db 248 KKALEIVRH-----SKDPLAKFQKQSDVDLYAIKEFFLNPPVT-DYNNLVWRDP 295
Qy 302 DEGLISFLVNDGFNEDRVTKAIEKISAKNKSOGRLSEFFK 345
Db 296 DEEGILKFLCDEHDFSEERVKNGLERLKAISGKOSTLESWFK 339

RESULT 14


```

; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-823-516-136

Query Match      33.4%; Score 648.5; DB 2; Length 340;
Best local Similarity 41.3%; Pred. No. 5.4e-56;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

Qy 19 KEQFFSYFGRKIAVDASMSIYQELIVVGTGMEITNEAGEVTSHLQGMFNRTIRLLEA 78
Db 12 KEIELENLYGKKIAIDALNAIYQELSTIRQDGTPLMDSKGRITSHLSGLEFVRTINLMEA 71

Qy 79 GIKPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKDALEKLSKRTVKVTRQH 138
Db 72 GIKPVVFDGPEPEPKKKELEKREAREAEKREKREALEKGELEARKYAOQATRVNEML 131

Qy 139 NEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASEDMDSLTGAPRFLHLM DP 198
Db 132 IEDAKKLELMGIPVQAPSGEQAAYMAAKGSVYASASQDDYDLSLLFGAPRLVRLTIT 191

Qy 199 SSKKIP-----VMEFDVAKVLELEITMDQFIDLCILCGDY-CDSIKGIGG 244
Db 192 GKKLPCKGNVVEIKPELTILE---EVLKELKLTREKLTIELAILVGTDYNFGGKIGL 247

Qy 245 QTALKLIRHGSGSTESILENLAKD---RYQIPEDWPYQEARLFKEPNVTLDIPELKWTPAP 301
Db 248 KKALETIVRH-----SKDPLAKFQKQSDVLYATKEFFLNPPVT-DNYNLVWRDP 295

Qy 302 DEGLISFLVKONGNEDRVTKATEKTKSAKNKSSOGRLSEFFK 345
Db 296 DEEGILKFLCDEHDFSEERVKNGLERLKAISKQKQSTLESWPK 339

```

Search completed: May 8, 2003, 21:44:30
Job time : 35 secs

QY 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 Db 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 QY 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 Db 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 QY 361 DKTSKAAANKTKAGGKK 379
 Db 361 DKTSKAAANKTKAGGKK 379

RESULT 2

US-09-805-311-6
 ; Sequence 6, Application US/09805311
 ; Patent No. US20010034886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mahajan, Pramod B.
 ; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
 ; FILE REFERENCE: 0961D
 ; CURRENT APPLICATION NUMBER: US/09/805,311
 ; CURRENT FILING DATE: 2001-03-13
 ; PRIOR APPLICATION NUMBER: 09/426,557
 ; PRIOR FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/112,332
 ; PRIOR FILING DATE: 1998-12-15
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-805-311-6

Query Match 100.0%; Score 1939; DB 10; Length 379;
 Best Local Similarity 100.0%; Pred. No. 2e-153;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGIGKLTLLADNAPKAMKEQKFESYFGKTIADVDSMSIYQFLIVVGRGTMTLTNEAGE 60
 Db 1 MGIGKLTLLADNAPKAMKEQKFESYFGKTIADVDSMSIYQFLIVVGRGTMTLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDDATKOLTEAVEVGD 120
 Db 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDDATKOLTEAVEVGD 120
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
 Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
 QY 181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCLGCGDYCDSIK 240
 Db 181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCLGCGDYCDSIK 240
 QY 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 Db 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 QY 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 Db 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 QY 361 DKTSKAAANKTKAGGKK 379
 Db 361 DKTSKAAANKTKAGGKK 379

RESULT 3

US-09-805-311-4
 ; Sequence 4, Application US/09805311

Patent No. US20010034886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mahajan, Pramod B.
 ; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
 ; FILE REFERENCE: 0961D
 ; CURRENT APPLICATION NUMBER: US/09/805,311
 ; CURRENT FILING DATE: 2001-03-13
 ; PRIOR APPLICATION NUMBER: 09/426,557
 ; PRIOR FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/112,332
 ; PRIOR FILING DATE: 1998-12-15
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-09-805-311-4

Query Match 99.7%; Score 1933; DB 10; Length 379;
 Best Local Similarity 99.7%; Pred. No. 6.2e-153;
 Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGIGKLTLLADNAPKAMKEQKFESYFGKTIADVDSMSIYQFLIVVGRGTMTLTNEAGE 60
 Db 1 MGIGKLTLLADNAPKAMKEQKFESYFGKTIADVDSMSIYQFLIVVGRGTMTLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDDATKOLTEAVEVGD 120
 Db 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDDATKOLTEAVEVGD 120
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
 Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
 QY 181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCLGCGDYCDSIK 240
 Db 181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCLGCGDYCDSIK 240
 QY 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 Db 241 GIGGOTALKLIHQSGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
 QY 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 Db 301 PDEGLISFLVKGNGFNEQVTKAEIKSAKNSQGRLESFFKPTATTAPLKRKETS 360
 QY 361 DKTSKAAANKTKAGGKK 379
 Db 361 DKTSKAAANKTKAGGKK 379

RESULT 4

US-09-805-311-8
 ; Sequence 8, Application US/09805311
 ; Patent No. US20010034886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mahajan, Pramod B.
 ; TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
 ; FILE REFERENCE: 0961D
 ; CURRENT APPLICATION NUMBER: US/09/805,311
 ; CURRENT FILING DATE: 2001-03-13
 ; PRIOR APPLICATION NUMBER: 09/426,557
 ; PRIOR FILING DATE: 1999-10-22
 ; PRIOR APPLICATION NUMBER: 60/112,332
 ; PRIOR FILING DATE: 1998-12-15
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 379
 ; TYPE: PRT

ORGANISM: Zea mays

US-09-805-311-8

Query Match 99.7%; Score 1933; DB 10; Length 379;

Best Local Similarity 99.7%; Pred No. 6.2e-155; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGKGLTKLLADNAPKAMKEQKPFSEYGRKIAVDASMSYQPLIVGRTGNETLTNEAGE 60
Db 1 MGKGLTKLLADNAPKAMKEQKPFSEYGRKIAVDASMSYQPLIVGRTGNETLTNEAGE 60
Qy 61 VTSHLQGMNRTIRLEAGIKPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG 120
Db 61 VTSHLQGMNRTIRLEAGIKPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG 120
Qy 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLHLMDFSSKKPIPMVEFDVAKVLEELTMDQFIDLCLCGGDCYCDSTK 240
Db 181 KDSLTFGAPRFLHLMDFSSKKPIPMVEFDVAKVLEELTMDQFIDLCLCGGDCYCDSTK 240
Qy 241 GIGGOTALKILIRHGSISILENENKDRYQIPEDWPYQEARLKPENVTLDIPELKWTA 300
Db 241 GIGGOTALKILIRHGSISILENENKDRYQIPEDWPYQEARLKPENVTLDIPELKWTA 300
Qy 301 PDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTISAPLKRKES 360
Db 301 PDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTISAPLKRKES 360
Qy 361 DKTSKAANKTKAGGKKK 379
Db 361 DKTSKAANKTKAGGKKK 379

RESULT 5

US-10-033-297-137
Sequence 137, Application US/10033297
Publication No. US20020187486A1

GENERAL INFORMATION:

APPLICANT: Hall, Jeff G.

Lyamichiev, Victor I.

Mast, Andrea L.

Brow, Mary Ann D.

TITLE OF INVENTION: Detection of Nucleic Acids By Multiple

Sequential Invasive Cleavages

NUMBER OF SEQUENCES: 163

CORRESPONDENCE ADDRESS:

ADDRESS: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States Of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/033,297

FILING DATE: 12-JAN-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/350,597

FILING DATE: 09-Jul-1999

APPLICATION NUMBER: US/08/823,516

FILING DATE: 24-MAR-1997

APPLICATION NUMBER: PCT/US97/01072

FILING DATE: 21-JAN-1997

APPLICATION NUMBER: US 08/759,038

FILING DATE: 02-DEC-1996

APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FORS-02736

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 137:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

STRANDEDNESS: No. US20020187486A1 Relevant

TOPOLOGY: No. US20020187486A1 Relevant

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 137:

US-10-033-297-137

Query Match 53.2%; Score 1032; DB 9; Length 380;

Best Local Similarity 53.8%; Pred. No. 5.8e-78;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

Qy 1 MGKGLTKLLADNAPKAMKEQKPFSEYGRKIAVDASMSYQPLIVGRTGNETLTNEAGE 60
Db 1 MGKGLTKLLADNAPKAMKEQKPFSEYGRKIAVDASMSYQPLIVGRTGNETLTNEAGE 60
Qy 61 VTSHLQGMNRTIRLEAGIKPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG 120
Db 61 VTSHLQGMNRTIRLEAGIKPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG 120
Qy 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLHLMDFSSKKPIPMVEFDVAKVLEELTMDQFIDLCLCGGDCYCDSTK 240
Db 181 MDSLTFGAPRFLHLMDFSSKKPIPMVEFDVAKVLEELTMDQFIDLCLCGGDCYCDSTK 240
Qy 241 GIGGOTALKILIRHGSISILENENKDRYQIPEDWPYQEARLKPENVTLDIPELKW 298
Db 241 GIGGOTALKILIRHGSISILENENKDRYQIPEDWPYQEARLKPENVTLDIPELKW 298
Qy 299 TAPDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTISAPLKRKE 358
Db 299 TAPDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTISAPLKRKE 358
Qy 359 TSDEK--TSKAANKTKAGGKKK 379
Db 359 TSDEK--TSKAANKTKAGGKKK 379

RESULT 6

US-09-940-244-137

Sequence 137, Application US/09940244

Publication No. US20030044796A1

GENERAL INFORMATION:

APPLICANT: Neri, Bruce P.

APPLICANT: Hall, Jeff G.

APPLICANT: Lyamichiev, Victor

APPLICANT: Smith, Lloyd M.

TITLE OF INVENTION: Reactions on Dendrimer

FILE REFERENCE: FORS-06478

CURRENT APPLICATION NUMBER: US/09/940,244

CURRENT FILING DATE: 2002-05-06

NUMBER OF SEQ ID NOS: 422

SOFTWARE: Patent in version 3.1

SEQ ID NO 137
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-940-244-137

Query Match 53.2%; Score 1032; DB 9; Length 380;

Best Local Similarity 53.8%; Pred. No. 5.8e-78;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGLTKLADNAPKAMEKQFESYGRKIAVDASMSIYQFLIIVGRTGNETLTNAGE 60
DB 1 MGIGLTKLADNAPKAMEKQFESYGRKIAVDASMSIYQFLIIVGRTGNETLTNAGE 60
QY 61 VTSHQGMFNRTRLEAGIKPVYVFDGKPPDKKQELAKRYKRDATKDLTAEVGD 120
DB 60 TTSILMGFVYRTIRMENGIRPVYVFDGKPPDKKQELAKRYKRDATKDLTAEVGD 120
QY 121 KDTEKLSKTVKTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
DB 120 EOEVEKTKRLVKTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
QY 181 MDSITFCAPRFLRLMDPSSKKIPVMEFDVAKVLELELTMDQFIDLICLGCDCYSIK 240
DB 180 MDCITFCSPVLMHRLTASPAKRLQIEFHLRLQELGNOFVLDLILGSDYCSIR 239
QY 241 GIGGOTALKILROHGSIESIENLKNDRYOIPEDWPYQBARLFKEPNVTL--IPELKW 298
DB 240 GIGPKRAVDLIQKHKSIEEIVRRDPNKNYPVNPENLHKEAHLFEPEV-LDPESVELKW 298
QY 299 TAPDEGLIFLVKNGFNEORVTKAEIKSAKNKSSQGRLESFKEPTATTAPLKRKE 358
DB 299 SEPNEELIKFMGEGKQFSEIRSGVKRLSKRQSTQGRDLDDFFKVTGSLSS-AKRKE 357
QY 359 TSDK--TSAKAAANKTKAGKKK 379
DB 358 PEPKSGSTKKA---KTGAAGKFK 377

RESULT 7

US-09-805-311-10
Sequence 10, Application US/09805311
Patent No. US20010034886A1
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961D
CURRENT APPLICATION NUMBER: US/09/805,311
PRIOR FILING DATE: 2001-03-13
PRIOR FILING DATE: 1999-10-22
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-805-311-10

Query Match 53.2%; Score 1032; DB 10; Length 380;

Best Local Similarity 53.8%; Pred. No. 5.8e-78;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGLTKLADNAPKAMEKQFESYGRKIAVDASMSIYQFLIIVGRTGNETLTNAGE 60
DB 1 MGIGLTKLADNAPKAMEKQFESYGRKIAVDASMSIYQFLIIVGRTGNETLTNAGE 60
QY 61 VTSHQGMFNRTRLEAGIKPVYVFDGKPPDKKQELAKRYKRDATKDLTAEVGD 120
DB 60 TTSILMGFVYRTIRMENGIRPVYVFDGKPPDKKQELAKRYKRDATKDLTAEVGD 120

QY 121 KDTEKLSKTVKTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 180
DB 120 EOEVEKTKRLVKTRQHNEDCKRLRLMGVVPVVEAPSEAECAALCINDKVFVASED 179
QY 181 MDSITFCAPRFLRLMDPSSKKIPVMEFDVAKVLELELTMDQFIDLICLGCDCYSIK 240
DB 180 MDCITFCSPVLMHRLTASPAKRLQIEFHLRLQELGNOFVLDLILGSDYCSIR 239
QY 241 GIGGOTALKILROHGSIESIENLKNDRYOIPEDWPYQBARLFKEPNVTL--IPELKW 298
DB 240 GIGPKRAVDLIQKHKSIEEIVRRDPNKNYPVNPENLHKEAHLFEPEV-LDPESVELKW 298
QY 299 TAPDEGLIFLVKNGFNEORVTKAEIKSAKNKSSQGRLESFKEPTATTAPLKRKE 358
DB 299 SEPNEELIKFMGEGKQFSEIRSGVKRLSKRQSTQGRDLDDFFKVTGSLSS-AKRKE 357
QY 359 TSDK--TSAKAAANKTKAGKKK 379
DB 358 PEPKSGSTKKA---KTGAAGKFK 377

RESULT 8

US-10-033-297-138
Sequence 138, Application US/10033297
Publication No. US20020187486A1

GENERAL INFORMATION:

APPLICANT: Hall, Jeff G.

Mast, Andrea L.

Brow, Mary Ann D.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

Lyamichev, Victor I.

TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020187486A1 Relevant
TOPOLOGY: No. US20020187486A1 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-10-033-297-138

Query Match 52.3%; Score 1014; DB 9; Length 378;
Best Local Similarity 53.2%; Pred. No. 1.8e-76;
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps

QY 1 MGIGKGLKLIADNAPKAMKEQFESYFGKRIADVADSMISYQFLIVVGRGTMTLTNEAGE 60
DB 1 MGIGHLAKLIADVAPSAIRENDIKSYFGKRVKVAIDASMSYQFLIAV-ROGGDVLQNEGE 59
QY 61 VTSHLOGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 60 TTS-LMGMYFRTIR-MENGKIPVYVFDGKPPOLKSGELAKRSERRAEKQLQQAQOAGM 117
QY 121 KDAIEKLSKRTVKTROHNEDCKRLRLMLGVVVEAPSEAECAALCINDKVFVASED 180
DB 118 EEEVEKTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALAKAGKVYAAATED 177
QY 181 MDSLTGAPRFLRHLMDPSSKKIPVMEFDVAKYLEELTMDQFIDLCILCGDCYDSIK 240
DB 178 MDCITFGSPVLMRHLTASEAKKLPQEFHLSRVQLQELNQEQFVDCILILGSDYCESIR 237
QY 241 GIGQATALKLIRHGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVT-LDIPELKWT 299
DB 238 GIGAKRAVDLIQKHKSIEEIVRLDPSKYVPYVPEWNLHKEAQQLFLEPEVVDPESELKWS 297
QY 300 APDEGLISFLVKNDFNEDRVTKAEIKKSAKNSSQGRLESEFFKPTATTAPLARKKET 359
DB 298 EPNEEELVKFMCGEKFSEERIRSGVKRLSKRSQGSTQGRLLDFFKVTGSLSS-AKRKEP 356
QY 360 SDKTSKAAANKTKAGGKK 379
DB 357 E---PKGPAKKAKTGGAGK 373

RESULT 9
US-09-940-244-138
Sequence 138, Application US/09940244
Publication No. US20030044796A1
GENERAL INFORMATION:
APPLICANT: Neri, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on Dendrimers
FILE REFERENCE: FORS-06478
CURRENT APPLICATION NUMBER: US/09/940,244
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: Patentin version 3.1
SEQ ID NO 138
LENGTH: 378
TYPE: PRT
ORGANISM: Mus musculus
US-09-940-244-138

Query Match 52.3%; Score 1014; DB 9; Length 378;
Best Local Similarity 53.2%; Pred. No. 1.8e-76;
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps

QY 1 MGIGKGLKLIADNAPKAMKEQFESYFGKRIADVADSMISYQFLIVVGRGTMTLTNEAGE 60
DB 1 MGIGHLAKLIADVAPSAIRENDIKSYFGKRVKVAIDASMSYQFLIAV-ROGGDVLQNEGE 59

QY 61 VTSHLOGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 60 TTS-LMGMYFRTIR-MENGKIPVYVFDGKPPOLKSGELAKRSERRAEKQLQQAQOAGM 117
QY 121 KDAIEKLSKRTVKTROHNEDCKRLRLMLGVVVEAPSEAECAALCINDKVFVASED 180
DB 118 EEEVEKTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALAKAGKVYAAATED 177
QY 181 MDSLTGAPRFLRHLMDPSSKKIPVMEFDVAKYLEELTMDQFIDLCILCGDCYDSIK 240
DB 178 MDCITFGSPVLMRHLTASEAKKLPQEFHLSRVQLQELNQEQFVDCILILGSDYCESIR 237
QY 241 GIGQATALKLIRHGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVT-LDIPELKWT 299
DB 238 GIGAKRAVDLIQKHKSIEEIVRLDPSKYVPYVPEWNLHKEAQQLFLEPEVVDPESELKWS 297
QY 300 APDEGLISFLVKNDFNEDRVTKAEIKKSAKNSSQGRLESEFFKPTATTAPLARKKET 359
DB 298 EPNEEELVKFMCGEKFSEERIRSGVKRLSKRSQGSTQGRLLDFFKVTGSLSS-AKRKEP 356
QY 360 SDKTSKAAANKTKAGGKK 379
DB 357 E---PKGPAKKAKTGGAGK 373

RESULT 10
US-10-033-297-139
Sequence 139, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Correspondence Address:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDNESS: NO. US20020187486A1 Relevant
TOPOLOGY: NO. US20020187486A1 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 139:
US-10-033-297-139

Query Match 48.4%; Score 938; DB 9; Length 382;
Best Local Similarity 49.5%; Pred. No. 3.8e-70;

Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMKQKESYFGRKIAVDASMSIYQFLIYVGRGTGNETITNEAGE 60
DB 1 MGKGLNIAIIEHVPSAIRKSDIKSFTGRKVAIDASMSLYQFLIYVGRGTGNETITNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 61 TTSHLMGMFYRTLRMDINGIKPCYVFDGKPPDLKSHLTKRSSRVETEKKLAEATTELE 120
QY 121 KDAIEKLSKRTVKVTRHNEDECKRLRLMGVVPVVEAPSEAEACALCINDKVFAYASED 180
DB 121 K--MQERRLVKVSKEHNEEAQKLLGLMGIPYIAPTAEAAQCAELAKKGVYAAASED 177
QY 181 MDSLTGAPRFLRHLMPPSSKKIPVMEFVAKVLELELTMDQFIDLCILGCDYCDISIK 240
DB 181 MDTLCYRTPFLRLHLTFSKAKKEPIHEIDTLVLRGLDTIYEQFVLDLCIMLCDCYCESIR 237
QY 241 GIGGOTALKLIHQHSIESILENL----NKDRYQIPEDWPYQOEARRLFKEPNVTLDIPE 295
DB 241 GIGGOTALKLIHQHSIESILENL----NKDRYQIPEDWPYQOEARRLFKEPNVTLDIPE 295
QY 296 --LKTAPDEGLISFLVKGNGFNEDEVRVTKAIEKISAKNKSQGRLESFFKPTATT--- 350
DB 296 --LKTAPDEGLISFLVKGNGFNEDEVRVTKAIEKISAKNKSQGRLESFFKPTATT--- 350
QY 351 -SAPLKRKETSOKTSAKAAANKTK 373
DB 357 LAAAKRAQENKLNK-NKNKVTK 379

RESULT 11
US-09-940-244-139

Sequence 139, Application US/09940244
Publication No. US20030044796A1
GENERAL INFORMATION:
APPLICANT: Neri, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on Dendrimers
FILE REFERENCE: FORS-06478
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: PatentIn version 3.1
SEQ ID NO 139
LENGTH: 382
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-940-244-139

Query Match 48.4%; Score 938; DB 9; Length 382;
Best Local Similarity 49.5%; Pred. No. 3.8e-70;

Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMKQKESYFGRKIAVDASMSIYQFLIYVGRGTGNETITNEAGE 60

DB 1 MGKGLNIAIIEHVPSAIRKSDIKSFTGRKVAIDASMSLYQFLIYVGRGTGNETITNEAGE 60
QY 61 VTSHLQGMFNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 61 TTSHLMGMFYRTLRMDINGIKPCYVFDGKPPDLKSHLTKRSSRVETEKKLAEATTELE 120
QY 121 KDAIEKLSKRTVKVTRHNEDECKRLRLMGVVPVVEAPSEAEACALCINDKVFAYASED 180
DB 121 K--MQERRLVKVSKEHNEEAQKLLGLMGIPYIAPTAEAAQCAELAKKGVYAAASED 177
QY 181 MDSLTGAPRFLRHLMPPSSKKIPVMEFVAKVLELELTMDQFIDLCILGCDYCDISIK 240
DB 181 MDTLCYRTPFLRLHLTFSKAKKEPIHEIDTLVLRGLDTIYEQFVLDLCIMLCDCYCESIR 237
QY 241 GIGGOTALKLIHQHSIESILENL----NKDRYQIPEDWPYQOEARRLFKEPNVTLDIPE 295
DB 241 GIGGOTALKLIHQHSIESILENL----NKDRYQIPEDWPYQOEARRLFKEPNVTLDIPE 295
QY 296 --LKTAPDEGLISFLVKGNGFNEDEVRVTKAIEKISAKNKSQGRLESFFKPTATT--- 350
DB 296 --LKTAPDEGLISFLVKGNGFNEDEVRVTKAIEKISAKNKSQGRLESFFKPTATT--- 350
QY 351 -SAPLKRKETSOKTSAKAAANKTK 373
DB 357 LAAAKRAQENKLNK-NKNKVTK 379

RESULT 12

US-10-033-297-79
Sequence 79, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 79:
US-10-033-297-79

Query Match 33.4%; Score 648.5; DB 9; Length 340;
Best Local Similarity 41.3%; Pred. No. 4e-46;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQKFSYGRKINDASMSYQFLIVVGRGTMETLTNEAGEVTSHLQGMFNRTIRLEA 78
Db 12 KELELENLGGKIAIDALNAIQFLSTIRQKDGTPLMDSKGRITSHLSGLFYRTINLMEA 71
QY 79 GIKPVYVFGKPPDMKQELAKRYSKRDDATKDLTEAVEVGGDKDAIEKLSKRTVYKTRQH 138
Db 72 GIKPVYVFGDEPPFKKELEKREAREAEAEKREALEKEGEIEARKYAQRAIVRNEML 131
QY 139 NEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASEDMDSLTFCGAPRFLHMDP 198
Db 132 IEDAKKLELGMGIPVQAPSEGAQAAYMAAGSVYASQDYDSLLFGAPRLVRLNLTIT 191
QY 199 SSKKIP-----VMEFDVAKVLEELTMDQFIDLCILCGDY-CDSIKGIGG 244
Db 192 GKRKLGKNNVVEIKPELIILE---EVLEKELTKREKUIELAILVGYDYNPGGKIGL 247
QY 245 OTALKLIROHGSIESILENLAKO---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWTP 301
Db 248 KKALEIVRH-----SKDPLAKFQKQSDVDLYAIKEFFLNPPTV-DNYLVWRDP 295
QY 302 DEGLISFLVKNQNGNEDRVTKAIEKIKSAKNSQGRLESFFK 345
Db 296 DEEGILKFLCDHDFSEERVKNGLERLKAISKGKSTLESWFK 339

RESULT 13
US-10-033-297-136
Sequence 136, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Brow, Mary Ann D.
Mast, Andrea L.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-No. US20020187486A1-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
SYNOPSIS: No. US20020187486A1 Relevant
TOPOLOGY: No. US20020187486A1 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-10-033-297-136

Query Match 33.4%; Score 648.5; DB 9; Length 340;
Best Local Similarity 41.3%; Pred. No. 4e-46;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQKFSYGRKIAVDASMSYQFLIVVGRGTMETLTNEAGEVTSHLQGMFNRTIRLEA 78
Db 12 KELELENLGGKIAIDALNAIQFLSTIRQKDGTPLMDSKGRITSHLSGLFYRTINLMEA 71
QY 79 GIKPVYVFGKPPDMKQELAKRYSKRDDATKDLTEAVEVGGDKDAIEKLSKRTVYKTRQH 138
Db 72 GIKPVYVFGDEPPFKKELEKREAREAEAEKREALEKEGEIEARKYAQRAIVRNEML 131
QY 139 NEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASEDMDSLTFCGAPRFLHMDP 198
Db 132 IEDAKKLELGMGIPVQAPSEGAQAAYMAAGSVYASQDYDSLLFGAPRLVRLNLTIT 191
QY 199 SSKKIP-----VMEFDVAKVLEELTMDQFIDLCILCGDY-CDSIKGIGG 244
Db 192 GKRKLGKNNVVEIKPELIILE---EVLEKELTKREKUIELAILVGYDYNPGGKIGL 247
QY 245 OTALKLIROHGSIESILENLAKO---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWTP 301
Db 248 KKALEIVRH-----SKDPLAKFQKQSDVDLYAIKEFFLNPPTV-DNYLVWRDP 295
QY 302 DEGLISFLVKNQNGNEDRVTKAIEKIKSAKNSQGRLESFFK 345
Db 296 DEEGILKFLCDHDFSEERVKNGLERLKAISKGKSTLESWFK 339

RESULT 14
US-09-940-244-79
Sequence 79, Application US/09940244
Publication No. US20030044796A1
GENERAL INFORMATION:
APPLICANT: Neri, Bruce P.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor
APPLICANT: Smith, Lloyd M.
TITLE OF INVENTION: Reactions on Dendrimers

FILE REFERENCE: FORS-06478
CURRENT APPLICATION NUMBER: US/09/940,244
CURRENT FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 422
SOFTWARE: PatentIn version 3.1
SEQ ID NO 79
LENGTH: 340
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-09-940-244-79

Query Match 33.4%; Score 648.5; DB 9; Length 340;

Best Local Similarity 41.3%; Pred. No. 4e-46;

Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRIADVADMSIYQFLVVGRTGMEITNAGEVTSHLQGMFNRTIRLEA 78

DB 12 KEIELENYGKTAIDALNAIQFLSTIRQDGTPLMDSKGRITSHLSGLFYRTINLMEA 71

QY 79 GIKPVYFDGKPPDMKKOELAKRYSKRDDATKDLTEAVEVGDKDAIEKLSKRTVKVTRQH 138

DB 72 GIKPVYFDGEPPEKKKEKREAREAEAEKWRKREALEKGBIEARKYAQRATRVNEML 131

QY 139 NEDCKRLRLMGVPVVEAPSAEACALCINDKVFVASEDMDSLTGAPRFLRLHMDP 198

DB 132 IEDAKKLELMGPIVQAPSEGEAAQAAVMAAGSVYASQDYDLSLLFGAPRLVRLNLTIT 191

QY 199 SSKKIP-----VMEFDVAKVLELELTMDQFIDLCILCGCDY-CDSIKGIGG 244

DB 192 GKRKLPGKNVYVEIKPELIIIE-----EVLKELKLTREKLELAILVGTDYNPGGKIGIGL 247

QY 245 QTALKLIRQHGSIESTILENLKND---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWTPAP 301

DB 248 KKALEIVRH-----SKDPLAKFQKSDVDLYAIKEFFLNPPVT-DNYNLVWRDP 295

QY 302 DEGLISFLYKNGFNEDRVTKAIEKIKSAKNKSQGRLESFFK 345

DB 296 DEEGILAFCLDHDHDFSEERVKNGLERLAKAIKSGKQSTLESWFK 339

RESULT 15

US-09-940-244-136

Sequence 136, Application US/09/940244

Publication No. US2003004796A1

GENERAL INFORMATION:

APPLICANT: Neri, Bruce P.

APPLICANT: Hall, Jeff G.

APPLICANT: Lyamichev, Victor

APPLICANT: Smith, Lloyd M.

TITLE OF INVENTION: Reactions on Dendrimers

FILE REFERENCE: FORS-06478

CURRENT APPLICATION NUMBER: US/09/940,244

CURRENT FILING DATE: 2002-05-06

NUMBER OF SEQ ID NOS: 422

SOFTWARE: PatentIn version 3.1

SEQ ID NO 136

LENGTH: 340

TYPE: PRT

ORGANISM: Pyrococcus furiosus

US-09-940-244-136

Query Match

Best Local Similarity 41.3%; Pred. No. 4e-46;

Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRIADVADMSIYQFLVVGRTGMEITNAGEVTSHLQGMFNRTIRLEA 78

DB 12 KEIELENYGKTAIDALNAIQFLSTIRQDGTPLMDSKGRITSHLSGLFYRTINLMEA 71

QY 79 GIKPVYFDGKPPDMKKOELAKRYSKRDDATKDLTEAVEVGDKDAIEKLSKRTVKVTRQH 138

DB 72 GIKPVYFDGEPPEKKKEKREAREAEKWRKREALEKGBIEARKYAQRATRVNEML 131

QY 139 NEDCKRLRLMGVPVVEAPSAEACALCINDKVFVASEDMDSLTGAPRFLRLHMDP 198

DB 132 IEDAKKLELMGPIVQAPSEGEAAQAAVMAAGSVYASQDYDLSLLFGAPRLVRLNLTIT 191

QY 199 SSKKIP-----VMEFDVAKVLELELTMDQFIDLCILCGCDY-CDSIKGIGG 244

DB 192 GKRKLPGKNVYVEIKPELIIIE-----EVLKELKLTREKLELAILVGTDYNPGGKIGIGL 247

QY 245 QTALKLIRQHGSIESTILENLKND---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWTPAP 301

DB 248 KKALEIVRH-----SKDPLAKFQKSDVDLYAIKEFFLNPPVT-DNYNLVWRDP 295

QY 302 DEGLISFLYKNGFNEDRVTKAIEKIKSAKNKSQGRLESFFK 345

DB 296 DEEGILAFCLDHDHDFSEERVKNGLERLAKAIKSGKQSTLESWFK 339

Search completed: May 8, 2003, 21:45:42

Job time : 59 secs

RESULT 2

A56531

DNA structure-specific endonuclease FEN1 [validated] - human

N:Alternate names: flap endonuclease 1; rad2 protein homolog

C:Species: Homo sapiens (man)

C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

C:Accession: A56531; B56054

R:Hiraoka, L.R.; Harrington, J.J.; Gerhard, D.S.; Lieber, M.R.; Hsieh, C.L.

Genomics 25, 220-225, 1995

A:Title: Sequence of human FEN-1, a structure-specific endonuclease and chromosomal local

A:Reference number: A56531; MUID:95293376; PMID:7774922

A:Accession: A56531

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-380 <HIR>

A:Cross-references: GB:L37374; NID:9642089; PIDN:AAA91331.1; PID:9642090

R:Murray, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.R.; Carr, A.M.

Mol. Cell. Biol. 14, 4878-4888, 1994

A:Title: Structural and functional conservation of the human homolog of the Schizosaccha

A:Reference number: A56054; MUID:94277093; PMID:8007985

A:Accession: B56054

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-380 <MUR>

A:Cross-references: GB:S70527

C:Genetics:

A:Gene: GDB:FEN1

A:Cross-references: GDB:451178; OMIM:600393

A:Map position: 11q12-11q12

C:Function:

A:Description: cleaves substrates with unannealed 5'-tails [validated, MUID:20209428]

A>Note: needed during Okazaki fragment processing and long patch base excision repair,

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: DNA repair; endonuclease

Query Match 53.2%; Score 1032; DB 2; Length 380;

Best Local Similarity 53.8%; Pred. No. 1,1e-61;

Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGIGKLTLLADNAPKAMKQKFEYFGKRIADVADSMISYQFLIVVGRGTMETLTNEAGE 60

DB 1 MGIOGLAKLIADVAPSAIRENDIKSFYFGKRIADVADSMISYQFLIAV-RGGSDVLQNEEGE 59

QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120

DB 60 TTSHLMGMFYRTIRMENGIKPVYVFDGKPPDLKSGELAKRSERRAEKQLOQAAGA 119

QY 121 KDAIEKLSKRTVKYTRQHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180

DB 120 EOEVEKFTRLVKYTRQHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 179

QY 181 MDSLTFGAPRFLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCDYCDISIK 240

DB 180 MDCLTFGAPVRLHMLTSEAKKLPQEFHLSRLQELGINSQFVDLCILGSDYCESIR 239

QY 241 GIGGOTALKLIRHGSISILENLKDRQIPEDWPYQEARLKPEPNVTLDD--IPELKW 298

DB 240 GIGPRAVDLLOKHSIEIVRRLDPNKYPVPMNLHKEAHQLFLEPEV-LDPESVELKW 298

QY 299 TAPDEGLISFLVKGDFNEDRVTKALEKISAKNKSQGRLESFEPKPTATTAPLAKKE 358

DB 299 SEPNEEELIKPWCCEKQFSERTISGVKRLSKRSQSGTGRLLDFFKVTGSLSS-ARKKE 357

QY 359 TSDK--TSKAAANKTKAGGKKK 379

DB 358 PEPKGSSTKKA---KTGAAGKFK 377

RESULT 3

A56054

DNA repair endonuclease rad2 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

C:Accession: A56054; T45210; T38725
R:Murray, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.R.; Carr, M.O. Cell. Biol. 14, 4878-4888, 1994

A:Title: Structural and functional conservation of the human homolog of the Schizosac

A:Reference number: A56054; MUID:94277093; PMID:8007985

A:Accession: A56054

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <MUR>

A:Cross-references: GB:S70605

A:Accession: T45210

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-380 <MUR>

A:Cross-references: EMBL:X77041; NID:g2909339; PIDN:CAB36991.1; PID:94456854

A:Experimental source: strain 972h(-)

R:Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A:Reference number: 221797

A:Accession: T38725

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-380 <GEN>

A:Cross-references: EMBL:Z99167; PIDN:CAB16282.1; GSPDB:GNO0066; SPDB:SPAC3G6.06c

A:Experimental source: strain 972h-; cosmid C366

C:Genetics:

A:Gene: rad2

A:Map position: 1

A:Introns: 5/1; 35/2

C:Function:

A:Description: DNA repair

C:Superfamily: conserved hypothetical protein YKL113c

Query Match 52.8%; Score 1023.5; DB 2; Length 380;

Best Local Similarity 50.7%; Pred. No. 3.9e-61;

Matches 194; Conservative 73; Mismatches 107; Indels 9; Gaps 3;

QY 1 MGIGKLTLLADNAPKAMKQKFEYFGKRIADVADSMISYQFLIVVGRGTMETLTNEAGE 60

DB 1 MGIGKLTLLADNAPKAMKQKFEYFGKRIADVADSMISYQFLIVVGRGTMETLTNEAGE 60

QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120

DB 61 TTSHLMGMFYRTIRLVNDGKIPCFVFDGKPPDLKSGELAKRVARHOKAREDOEE*KEVGT 120

QY 121 KDAIEKLSKRTVKYTRQHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180

DB 121 AEMVDFAKRTVKYTRQHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180

QY 181 MDSLTFGAPRFLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILGCDYCDISIK 240

DB 181 MDCLTFGAPVRLHMLTSEAKKLPQEFHLSRLQELGINSQFVDLCILGSDYCESIR 240

QY 241 GIGGOTALKLIRHGSISILENLKDRQIPEDWPYQEARLKPEPNVTLDDIP----EL 296

DB 241 GIGGOTALKLIRHGSISILENLKDRQIPEDWPYQEARLKPEPNVTLDDIP----EL 296

QY 297 TAPDEGLISFLVKGDFNEDRVTKALEKISAKNKSQGRLESFEPKPTATTAPLAKR 356

DB 297 TAPDEGLISFLVKGDFNEDRVTKALEKISAKNKSQGRLESFEPKPTATTAPLAKR 356

QY 357 KETSDEKTSKAAANKTKAGGKKK 379

DB 356 VDTKSKGSAKRKRDNSNKGESKK 378

RESULT 4

A53730

flap endonuclease 1 - mouse

N:Alternate names: FEN-1

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999

C:Accession: A53730

R:Harrington, J.J.; Lieber, M.R.

Genes Dev 8, 1344-1355, 1994

A:Title: Functional domains within FEN-1 and RAD2 define a family of structure-specific

A:Reference number: A53730; MUID:95011546; PMID:7926735

A:Accession: A53730

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-378 <HAR>

A:Cross-references: GB:L26320; NID:g499648; PIDN:AAC37664.1; PID:g499649

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: DNA repair; endonuclease

Query Match 52.3%; Score 1014; DB 2; Length 378;
Best Local Similarity 53.2%; Pred. No. 1.7e-60;
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGIGLTKLLADNAPKAMKEQKFSYGRKIYAVDASMSIYQFLIVVGRGTMTLTNEAGE 60

DB 1 MGIIHLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIIV-ROGGDVLTNEAGE 59

QY 61 VTSHLOGMFRNTRIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATFDLTEAVEVGD 120

DB 60 TTS-LMGMFYTRIR-WENGIRKPVYVFDGKPPDLKSGELAKRAREAEKQLOQAEAGM 117

QY 121 KDTEKLSKRTVKVTRQHNEDCKRLRLMGMVVPVEAPSEAEACALCINDKVFVAVASED 180

DB 118 EEEVEKFTKRLVKTQHNEDCKHLLSLMGIPIYLDAPSEAEACALAKAGKVVYAAATED 177

QY 181 MDSLTGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCTLCGCDYCSIR 240

DB 178 MDCLTFGSPVLMRHLTASEAKKLTPQEPHLSRVLQELGLNQEFVLDCLILLGSDYCSIR 237

QY 241 GIGGOTALKLIROHGSIESILNKNDRYQIPEDWPYQEARRLFKPEPNT-LDIPELKWT 299

DB 238 GIGAKRAVDLQKHISIEIVRLDPSKYPVPENLWKEAQQFLPEVVDPESEVELWS 297

QY 300 APDEGLISPLVKNDFNEDRVTKAEIKSAKNSQGRLESFFK-----PTATTSAPL 354

DB 298 EPNEELVPMGCKEKFSEIRSRVSKLSKSGSTQGRDLDFFKVTGSLSS-ARRKEP 356

QY 360 SDKTSKAAANKTKAGGKK 379

DB 357 E---PKGPAKKAKTGGAGK 373

RESULT 5

Tl3692

Hypothetical protein EG0003.3 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: Tl3692

R:Murphy, L.; Harris, D.; Barrell, B.

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17699

A:Accession: Tl3692

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 <MUR>

A:Cross-references: EMBL:AL031863; PIDN:CAA21320.1

C:Genetics:

A:Introns: 28/1

A:Note: EG:EG0003.3

C:Superfamily: conserved hypothetical protein YKL113c

Query Match 51.5%; Score 998.5; DB 2; Length 385;
Best Local Similarity 52.9%; Pred. No. 1.9e-59;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;

QY 1 MGIGLTKLLADNAPKAMKEQKFSYGRKIYAVDASMSIYQFLIVVGRGTMTLTNEAGE 60

DB 1 MGIIHLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIIV-ROGGDVLTNEAGE 59

QY 61 VTSHLOGMFRNTRIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATFDLTEAVEVGD 120

DB 60 PTSHLMGMFYTRIRLDNDGIRKPVYVFDGKPPDLKSGELAKRAREAEKALKAATDAGD 119

QY 121 KDTEKLSKRTVKVTRQHNEDCKRLRLMGMVVPVEAPSEAEACALCINDKVFVAVASED 180

DB 120 DAGIEKFNRLRVTKHEAKAKELLTLMGVYVDAPCEAEACALAKAGKVVATATED 179

QY 181 MDSLTGAPRFLRLHMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCTLCGCDYCSIR 240

DB 180 MDALTGSKLRLYLYSEARKMPVKESEYDKLLEGLAINNREFIDLCTLCGCDYCSIR 239

QY 241 GIGGOTALKLIROHGSIESILNKNDRYQIPEDWPYQEARRLFKPEPNT-LDIPELKWT 299

DB 240 GIGKRAIELINTYRDIEFILNDSSKTYVPENWTKVARELFIEPEVADADSIDLKVV 299

QY 300 APDEGLISPLVKNDFNEDRVTKAEIKSAKNSQGRLESFFK-----PTATTSAPL 354

DB 300 EPDEGLVFLCGDRQFNEERVNGAKLMKSKQAQTOVRLDSFFKTLSTPNATNAA-- 357

QY 355 KRKETSCKTSAKAAANKTKAGG 376

DB 358 KKK--AEAAKSNANKKARTSG 377

RESULT 6

S22267

Hypothetical protein YKL113c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YKL510; RAD2 protein homolog

C:Species: Saccharomyces cerevisiae

C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 24-Sep-1999

C:Accession: S22267; S37941

R:Jacquier, A.; Legrain, P.; Dujon, B.

Yeast 8, 121-132, 1992

A:Title: Sequence of a 10.7 kb segment of yeast chromosome XI identifies the APN1 an

S.

A:Reference number: S22267; MUID:92221689; PMID:1561835

A:Accession: S22267

A:Molecule type: DNA

A:Residues: 1-382 <JAC>

A:Cross-references: GB:S93804; NID:g248391; PIDN:AAB21998.1; PID:g248392

A:Experimental source: strain S288C

R:Jacquier, A.; Legrain, P.; Collea, L.; Richard, G.F.; Thierry, A.; Dujon, B.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37938

A:Accession: S37941

A:Molecule type: DNA

A:Residues: 1-382 <JAC>

A:Cross-references: EMBL:Z28113; NID:g486189; PIDN:CAA81953.1; PID:g486190; MIPS:YKL1

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:RAD27

A:Cross-references: SGD:S0001596; MIPS:YKL113c

A:Map position: 11L

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: nucleus

Query Match 48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 2.1e-55;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGIGLTKLLADNAPKAMKEQKFSYGRKIYAVDASMSIYQFLIVVGRGTMTLTNEAGE 60

DB 1 MGIIHLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIIV-ROGGDVLTNEAGE 60

QY 61 VTSHLOGMFRNTRIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKOLTEAVEVGD 120

DB 61 TTSHLMGMFYTRIRLDNDGIRKPVYVFDGKPPDLKSGELAKRAREAEKALKAATTELE 120

QY 121 KDTEKLSKRTVKVTRQHNEDCKRLRLMGMVVPVEAPSEAEACALCINDKVFVAVASED 180

DB 121 K---MKQERLKVKSKEHNEAQLLGLMGIPYIATPEAEACALCINDKVFVAVASED 177

A:Cross-references: GB:AE001087; GB:AE000782; NID:g2689410; PIDN:AA890967.1; PID:g265037
C:Superfamily: conserved hypothetical protein YKL113c

Query Match 27.8%; Score 538.5; DB 2; Length 336;
Best Local Similarity 36.7%; Pred. No. 9.3e-29;
Matches 125; Conservative 67; Mismatches 118; Indels 31; Gaps 7;

QY 19 KEOKPESYGRKIAVDASMSYQQLVVGRTGMEITNEAGEVTSFLOGMENVRTIRLLEA 78
DB 12 EEVELEYSGKIIAIDAENTLYQISIRQDPGTPLKDSGRITSHLSILLYRVSNWVEV 71

QY 79 GKIPVYVFGKPPDMKKQELAKRYSKRDDATKOLTEAVEVGDKDAIEKLSKRTVKVTRQH 138
DB 72 GIRPVVFGEPEPFKAEIERKKRAEAEEMWIAALQAGKDA-KKYAQAAGRVEYI 130

QY 139 NEDCKRLRLMGVVPVVEAPSAEAEACALCINDKVFVASEDMDSITFGAPFLRLHMDP 198
DB 131 VDSAKTLLSYMGIPFVDAPSEGAQAAYMAAGDVETGSDYDSLLFGSPRLARNLAIT 190

QY 199 SSKKIP-----VMEFDVAKVLELELTMDQFDICILGCDYCDISIKIGGQ 245
DB 191 GKRKLPGKNVYVDVPEIIILESN-----LKRGLTREQLDIDIALVGTQYNEGVKGVK 246

QY 246 TALKLIRHOGSIESILENL--NKORYQIPEDWPYQEARRLFKEPNVTLDIPELKWTAPE 303
DB 247 KALNYIKTKGDIFFALKALKVNIDH-----VEETRNFFLNPVTTDDY-RIEFREPDF 297

QY 304 EGLISFLVKGDFNDRVTKAEIKSAKNSOGRLSEFF 344
DB 298 EKAIEFLCEHDFSRERKEALKAL--KSTQATLERWF 336

RESULT 13
H72765
probable flap endonuclease-1 APE0115 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H72765
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hailkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawa, H.; 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72765
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <KAW>
A:Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79026.1; PID:g5103505
A:Gene: APE0115
A:Experimental source: strain K1
C:Genetics:
C:Superfamily: conserved hypothetical protein YKL113c

Query Match 26.5%; Score 514.5; DB 2; Length 401;
Best Local Similarity 34.6%; Pred. No. 4.7e-27;
Matches 119; Conservative 74; Mismatches 132; Indels 19; Gaps 6;

QY 15 PKAMKEQKFESEYFGKRIAVDASMSYQQLVVGRTGMEITNEAGEVTSFLOGMENVRTIR 74
DB 62 PEAREVELRSLGVLGVALDAYNMLYQFLTAIRQPDGTPLLDREGRVTSFSLGSLFYRTIN 121

QY 75 LLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKOLTEAVEVGDKDAIEKLSKRTVKV 134
DB 122 LVEEGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKOLTEAVEVGDKDAIEKLSKRTVKV 181

QY 135 TRHNECKRLRLMGVVPVVEAPSAEAEACALCINDKVFVASEDMDSITFGAPFLRLH 194
DB 182 TSDMVEESKELLDAMGMPWQAPAGEAQAAYMARKDADATWQSDYDSLLFGSPRLVKN 241

QY 195 LMDPSSKKIPVME-----FDVAKVLELELTMDQFDICILGCDYCDISIKIGG 244
DB 242 LAITGRKLPGRDQYVEIKPEITIEPLLSKLTITREQLIANGILLGTDYNPGGVGRGYGP 301

QY 245 QTALKLIRHOGSIRSILENLNKORYQIPEDWPY-QEARRLFKEPNVTLDIPELKWTAPE 303
DB 302 KITALHLYSLGDPMKVVLASVPRGEY-----DPDLRLKRYVEYFLNPPVTTDDY-KIEFRKPDQ 356

QY 304 EGLISFLVKGDFNDRVTKAEIKSA-----KNKSOGRLSEFF 344
DB 357 DKVREILVERHDENPERVERALERLKGAYREKLRGQSRLDWMF 400

RESULT 14
F90158
DNA repair endo/exonuclease FEN-1 (RAD2) (rad2) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001;
C:Accession: F90158
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, J.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.; submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90158
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <KUR>
A:Cross-references: GB:AE006641; NID:gl3813313; PIDN:AAK40525.1; GSPDB:GN00155
C:Genetics:
A:Gene: rad2
C:Superfamily: conserved hypothetical protein YKL113c

Query Match 26.1%; Score 506; DB 2; Length 302;
Best Local Similarity 37.9%; Pred. No. 1.2e-26;
Matches 118; Conservative 58; Mismatches 97; Indels 38; Gaps 6;

QY 59 GVTSHLQGMNRTIRLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEV 118
DB 5 GVTSHLSGLFYRTINILEGVPIYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEV 64

QY 119 GKDIALEKLSKRTVKVTRHNECKRLRLMGVVPVVEAPSAEAEACALCINDKVFVAVAS 178
DB 65 GKIEELRYSQLRLSLNINVEESKLLRAMGPIYVQAPSEGEAAYLNKLGLSWAAS 124

QY 179 EDMDSLTGAPFLRLHMDPSKKIP-----VMEFDVAKVLELELTMDQFDIDL 227
DB 125 QYDAILFGAKRLVRLNLTITGKRLPNKVDYVEIKPEITETI--LLKLGITR9QLIDI 182

QY 228 CILGCDY-CDSIKGIGGOTAKLIRHOGSIESILE-----NLNKDRYQIPEDWPY 277
DB 183 GILIGTDYNDGIRGIGPERALKIYKIKEMVEGSELSKKDINFID----- 232

QY 278 QEARRLFKEPNVTLDIPELKWTAPEDEGLISFLVKGDFNDRVTKAEIK-----IKSAKN 333
DB 233 -ETRGLFNPQVVPPEALDILNEPNEGDIINILVYEHNFSEERVKNIGIERLTKAEARG 291

QY 334 KSSQGRLESEFF 344
DB 292 ASRQTCLDRWF 302

RESULT 15
F84290
DNA repair protein [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: F84290
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leibauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaung, K.H.; Alam, M.; Freitas, P.; Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M., A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84290

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:02:17 ; Search time 35 seconds

(without alignments)
449.129 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGIKGLTKLLADNAPKAKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1032	53.2	380	1	FENL_HUMAN	P39748	homo sapien
2	1023.5	52.8	380	1	RAD2_SCHPO	P39750	schizosacch
3	1014	52.3	378	1	FENL_MOUSE	P39749	mus musculus
4	938	48.4	382	1	RA27_YEAST	P26793	saccharomyc
5	560.5	28.9	326	1	FEN_METJA	Q58839	methanococc
6	260.5	13.4	571	1	EX01_SCHPO	P33695	schizosacch
7	245.5	12.7	496	1	YA31_SCHPO	Q09708	schizosacch
8	232.5	12.0	1031	1	RAD2_YEAST	P07276	saccharomyc
9	220.5	11.4	1186	1	XPG_HUMAN	P28715	homo sapien
10	220	11.3	1170	1	XPG_MOUSE	P35689	mus musculus
11	218	11.2	702	1	DIN7_YEAST	P39875	saccharomyc
12	204.5	10.5	430	1	DIN7_YEAST	Q12086	saccharomyc
13	204.5	10.5	1112	1	RAL3_SCHPO	P28706	schizosacch
14	200	10.3	1196	1	XPG_YENLA	P14629	xenopus lae
15	184.5	9.5	834	1	DPOL_THECA	P80194	thermus cal
16	176.5	9.1	886	1	DPOL_SYNT3	Q53971	synechocyst
17	173	8.9	832	1	DPOL_THERAQ	P19821	thermus aqu
18	172.5	8.9	834	1	DPOL_THERAQ	P52028	thermus the
19	166	8.6	289	1	EX53_AQUAE	Q67550	aquifex ae
20	157.5	8.1	831	1	DPOL_THERH	P30313	thermus the
21	156.5	8.1	833	1	DPOL_THERH	O52225	thermus fil
22	154.5	8.0	877	1	DPOL_STREN	P13252	streptococc
23	153.5	7.9	877	1	DPOL_LACLC	Q32801	lactococcus
24	151.5	7.8	759	1	YENL_YEAST	P40028	saccharomyc
25	149	7.7	930	1	DPOL_HAEIN	P43741	haemophilus
26	144	7.4	877	1	DPOL_BACCA	Q04957	bacillus ca
27	142.5	7.3	877	1	DPOL_LACIA	Q9cds1	lactococcus
28	138	7.1	921	1	DPOL_RICHE	Q9rib6	rickettsia
29	135.5	7.0	880	1	DPOL_BACSV	Q34996	bacillus su
30	134.5	6.9	956	1	DPOL_DEIRA	P52027	deinococcus
31	134	6.9	867	1	DPOL_RICPR	O05949	rickettsia
32	134	6.9	922	1	DPOL_RICPE	Q9raa9	rickettsia
33	123	6.3	876	1	DPOL_BACST	P52026	bacillus st

34	118.5	6.1	850	1	DPOL_ANATH	O50156	anaerocellu
35	113.5	5.9	942	1	DPOL_CHLAU	O08307	chloroflexu
36	112	5.8	317	1	ACCO_PEA	P31239	pisum sativ
37	112	5.8	984	1	SECA_AQUAE	C67718	aquifex ae
38	109.5	5.6	891	1	DPOL_HELPY	P56105	helicobacte
39	109.5	5.6	1016	1	DPOL_RHILE	Q98102	rhizobium l
40	109.5	5.6	1650	1	BP28_CAREL	Q23495	caenorhabdi
41	109	5.6	886	1	RA50_ARCFU	O29230	archaeoglob
42	109	5.6	897	1	DPOL_HELPFJ	Q921e9	helicobacte
43	108.5	5.6	444	1	TIG_CAMJE	Q46108	campylobacte
44	108.5	5.6	928	1	DPOL_ECOLI	P00582	escherichia
45	108.5	5.6	1509	1	MYSN_ACACA	P05659	acanthamoeb

ALIGNMENTS

RESULT 1
FENL_HUMAN
AC P39748; STANDARD; PRT; 380 AA.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE FLAP endonuclease-1 (Maturation factor 1) (MFI).
GN FENL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94277093; PubMed=8007985;
RA MURRAY J.M., TAVASSOLI M., AL-HARITHY R., SHELDRIK K.S.,
RA LEHMAN A.R., GARR A.M., WATTS F.Z.,
RT "Structural and functional conservation of the human homolog of the
RT Schizosaccharomyces pombe rad2 gene, which is required for chromosome
RT segregation and recovery from DNA damage";
RL Mol. Cell. Biol. 14:4878-4888(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RA LIEBER M.R., Hsieh C.L., HARRINGTON J.J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA LAMERDIN J.E., MCCREADY P.M., COLEMAN M., SKOWRONSKI E., ADAMSON A.W.,
RA BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STILWAGEN S.,
RA PHAN H., VELASCO N., DO L., REGALA W., TERRY A., GARNES J.,
RA DANGHAN L., POUNDSTONE P., CHRISTENSEN M., GEORGESCU A., AVILA J.,
RA LIU S., ATTIX C., ANDREISE T., TRANKHELM M., AMICO-KELLER G.,
RA COFIELD J., DUARTE S., LUCAS S., BRUCE R., THOMAS P., QUAN G.,
RA KRONMILLER B., ARELLANO A., MONTGOMERY M., OW D., NOLAN M., TRONG S.,
RA KOBAYASHI A., OLSEN A.S.,
RA CARRARO A.V.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA STRAUSBERG R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=95050647; PubMed=7961795;
RA ROBINS P., PAPPIN D.J., WOOD R.D., LINDAHL T.;
RT "Structural and functional homology between mammalian DNase IV and
RT the 5'-nuclease domain of Escherichia coli DNA polymerase I";
RL J. Biol. Chem. 269:28535-28538(1994).
CC -!- FUNCTION: ENDONUCLEASE THAT CLEAVE 5'FLAP STRUCTURE AND FAILS
CC TO CLEAVE OTHER DNA STRUCTURES, INCLUDING 3'FLAPS AND SINGLE
CC STRANDED DNA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FENL


```

CC CC SUBFAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; X76771; CAA54166.1; -
CC CC EMBL; L37374; AAA91331.1; -
CC CC EMBL; AC004770; AAC23394.1; -
CC CC EMBL; BC000323; AAH00323.1; -
CC CC HSSP; Q58839; 1A76.
CC CC Genew; HGNC:3650; FEN1.
CC CC MIM; 600393; -
CC CC InterPro; IPR002421; 5.3 exonuclease.
CC CC InterPro; IPR000513; Exo_N.1.
CC CC InterPro; IPR003584; HHH_2.
CC CC InterPro; IPR001532; XPGC_Rad.
CC CC Pfam; PF00752; XPG_N; 1.
CC CC Pfam; PF00867; XPG_I; 1.
CC CC PRINTS; PR00853; XPGRADSUPER.
CC CC SMART; SM00279; Hh2; 1.
CC CC SMART; SM00484; XPGI; 1.
CC CC SMART; SM00485; XPGN; 1.
CC CC PROSITE; PS00841; XPG_1; 1.
CC CC PROSITE; PS00842; XPG_2; 1.
CC CC Hydrolase; Nuclease; Endonuclease; Nuclear protein.
CC CC DOMAIN 1 104 N-DOMAIN.
CC CC FT DOMAIN 122 253 I-DOMAIN.
CC CC SQ SEQUENCE 380 AA; 42593 MW; 5154F2F6E57592C5 CRC64;

Query Match 53.28; Score 1032; DB 1; Length 380;
Best Local Similarity 53.88; Pred No. 2.2e-60;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMKEOKFESTFGKRIADVDSMSYQPLIVVGTGMEITNAGE 60
DB 1 MGIGQLKLIADVAPSATRENDKSYFGKVAIDASNSIYQELIIV-RGGGDVLONEEGE 59

QY 61 VTSHLQGFNRTIRLLEAGKPVYVFDGKPPDMKKOELAKRYSKRDDATKDLTFAVEVGD 120
DB 60 TTSHLGMNFYRTIRMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKOLQQAAGA 119

QY 121 KDAIEKLSKRTVKTQHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFASVED 180
DB 120 EQEVEKFKLVYKVKQHNDECKHLLSLMGIPYLDAPSEAEASCAALYKAGKVYAATED 179

QY 181 MSLTFGAPRFLHMDPSKKIPVMEFDYAKVLELELTMDQFTDGLCGDCYDSIK 240
DB 180 MDCLTFGSPVLMRHLTASEAKPLQEFHLSRLIQELGNQEFVDCLLGLSDYCESIR 239

QY 241 GIGGQALKLIRHGSTESILENKORYOIPEDNPYQEARLKFEPNVLID--IPELKW 298
DB 240 GIGPRKAVDLIQHKHSIEIIVRLDPNKNYPVNPENWLHKAHQLFEPEV-LDPESVELKW 298

QY 299 TAPDEGLTSLFVKDNGFNDRVTKAJEIKSAKNKSSQGRLESFETKPTATTSAPLKRKE 358
DB 299 SEFNEELIKFWGKEQFSEERISRGVKSRSQGSTQGRLDGDFKVTGSLSS-AKRKE 357

QY 359 TSDK--TSKAANKTKAGKKK 379
DB 358 PEPKGTGKKKA---KIGAAGFK 377

RESULT 2
RAD2_SCHPO STANDARD; PRT; 380 AA.
ID RAD2_SCHPO
AC P39750;
DT 01-FEB-1995 (Rel. 31, Created)

```

```

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA repair protein rad2.
GN RAD2 OR SPAC306.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RN [2] SEQUENCE FROM N.A.
RP MEDLINE=94277093; PubMed=8007985;
RX Murray J.M., Tavassoli M., Al-Harithy R., Sheldrick K.S.,
RA Lehmann A.R., Carr A.M., Watts F.Z.;
RT "Structural and functional conservation of the human homolog of the
RT Schizosaccharomyces pombe rad2 gene, which is required for chromosome
RT segregation and recovery from DNA damage.";
RL Mol. Cell. Biol. 14:4878-4888(1994).
RN [2]
RN [3] SEQUENCE FROM N.A.
RP MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Chillingworth T., Churcher C.M.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Woldjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Lehter E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Holzach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: REQUIRED FOR THE FIDELITY OF CHROMOSOME SEPARATION AT
CC MITOSIS AND WHICH IS ALSO INVOLVED IN THE RESPONSE TO DNA DAMAGE.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; X77041; CAB36991.1; -
CC CC EMBL; Z99167; CAB16282.1; -
CC CC HSSP; Q58839; 1A76.
CC CC InterPro; IPR000513; Exo_N.1.
CC CC InterPro; IPR003584; HHH_2.
CC CC InterPro; IPR001532; XPGC_Rad.
CC CC Pfam; PF00752; XPG_N; 1.
CC CC Pfam; PF00867; XPG_I; 1.
CC CC PRINTS; PR00853; XPGRADSUPER.
CC CC SMART; SM00279; Hh2; 1.
CC CC SMART; SM00484; XPGI; 1.

```

DR SMART: SM00485; XPGN: 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW DNA repair; Hydroxylase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 105 N-DOMAIN.
 FT DOMAIN 123 254 I-DOMAIN.
 SQ SEQUENCE 380 AA; 42866 MW; 3604065D9B34CBF CRC64;

Query Match 52.8%; Score 1023.5; DB 1; Length 380;
 Best Local Similarity 50.7%; Pred. No. 8e-60;
 Matches 194; Conservative 73; Mismatches 107; Indels 9; Gaps 3;

QY 1 MGIGLTKLLADNAPKAMKEQKESYFCRKIAVDASMSIYQFLIVVGRGTGNETLNEAGE 60
 DB 1 MGIGLAQVLSERAPASVKNHDKNIFGKRVDAIDASMSLYQFLIVRSODGQOLNEAGE 60
 QY 61 VTSHLOGMENRTIRLEAGIKPVYVDFGKPPDMKQELAKRYSKRDDATKDLTEAVEVGD 120
 DB 61 TTSHLGMFYRTIRVDNGIKFCFVDFGKPPDLKSGELAKRYAHQKAREDEEETKEVGT 120
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVVPEAPSEAEACALCINDKVFVASED 180
 DB 121 AEMVDFRAKTKVAVTRQHNEDCKRLRLMLGVVPEAPSEAEACALCINDKVFVASED 180
 QY 181 MDSITFGAPRFLRHLMDSPPSKKIPVMEFDVAKLVLEELTMDQFIDLCLCGCDYCSIK 240
 DB 181 MDITLGFQAPVLLRHLLTFSQKREPISEYNEIKALNGLDMSVEQFVLDCLLGLCYCEPIR 240
 QY 241 GIGGQTKALKIROHGSISILENLANKDRIQIPEDWPYQEARLKFEPNVTLDIP---EL 296
 DB 241 GVGPARAVELIRGYTLDRFVKEADRSKYPIDEDWPYEDARRFLDAEV---LPGEETEL 297
 QY 297 KWTAPDEGLISFLVKGDNFNDRTVTKAIEKIKSAKNKSSQGRLESFFKPTATTAPLKR 356
 DB 298 KWKSPDADGIIQLVKEKGFENDRVKLGINLEKASITIPQGRDLSFFRP--VPSPKPK 355
 QY 357 KETSQTSKAAANKTKAGGKK 379
 DB 356 VDTKSGAKRKDRSNKGESKK 378

RESULT 3
 ID FEN1_MOUSE STANDARD; PRT; 378 AA.
 AC P39749;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE FLAP endonuclease-1.
 OS FEN1 OR FEN-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 260-275 AND 331-352.
 RC STRAIN-BALB/c;
 RX MEDLINE=95011546; PubMed=7926735;
 RA Harrington J.J., Lieber M.R.;
 RT "Functional domains within FEN-1 and RAD2 define a family of
 RT structure-specific endonucleases: Implications for nucleotide excision
 RT repair.";
 RL Genes Dev. 8:1344-1355(1994).
 CC -!- FUNCTION: ENDONUCLEASE THAT CLEAVE 5' FLAP STRUCTURE AND FAILS
 CC TO CLEAVE OTHER DNA STRUCTURES, INCLUDING 3' FLAPS AND SINGLE
 CC STRANDED DNA.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
 CC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L26320; AAC37864.1; -;
 DR HSP; Q58839; IAY6.
 DR MGD; MGI:102779; Fen1.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N.1.
 DR InterPro; IPR003584; HHH 2.
 DR InterPro; IPR001532; XPGC_Rad.
 DR Pfam; PF00752; XPG_N; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR PRINTS; PR00853; XPGPADSUPER.
 DR SMART; SM00279; HHH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW Hydroxylase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 102 N-DOMAIN.
 FT DOMAIN 120 251 I-DOMAIN.
 SQ SEQUENCE 378 AA; 42314 MW; 827946BD8BDECF39 CRC64;

Query Match 52.3%; Score 1014; DB 1; Length 378;
 Best Local Similarity 53.2%; Pred. No. 3.3e-59;
 Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGIGLTKLLADNAPKAMKEQKESYFCRKIAVDASMSIYQFLIVVGRGTGNETLNEAGE 60
 DB 1 MGIGLAQVLSERAPASVKNHDKNIFGKRVDAIDASMSIYQFLIV-RGGDVLENEGE 59
 QY 61 VTSHLOGMENRTIRLEAGIKPVYVDFGKPPDMKQELAKRYSKRDDATKDLTEAVEVGD 120
 DB 60 TTS-LMGMYRTIR-MENGIRKPVYVDFGKPPDLKSGELAKRYAHQKAREDEEETKEVGT 117
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVVPEAPSEAEACALCINDKVFVASED 180
 DB 118 EEEVEKTRKLVKTRQHNEDCKRLRLMLGVVPEAPSEAEACALCINDKVFVASED 177
 QY 181 MDSITFGAPRFLRHLMDSPPSKKIPVMEFDVAKLVLEELTMDQFIDLCLCGCDYCSIK 240
 DB 178 MDCLTFGSPVLMRHLTASEAKKLPIDQEPHLSRVQLGELNQEQFVLDCLLGLSDYCESIR 237
 QY 241 GIGGQTKALKIROHGSISILENLANKDRIQIPEDWPYQEARLKFEPNVT-LDIPELKWT 299
 DB 238 GIGAKRAVDLIQKHKSIEIVRRLDPSKYPVPEWNLHKEAQQLLEPEVVDPESELKWS 297
 QY 300 APDEGLISFLVKGDNFNDRTVTKAIEKIKSAKNKSSQGRLESFFKPTATTAPLKR 359
 DB 298 EPNEEELVPMCEGKQFSEERIRSGVKRLSKSGROGTOGRDLDDFFKVTGSLSS-AKRKEP 356
 QY 360 SDKTSKAAANKTKAGGKK 379
 DB 357 E---PKGPAKKAKTGGAGK 373

RESULT 4
 ID RA27_YEAST STANDARD; PRT; 382 AA.
 AC P25703;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Structure specific endonuclease RAD27.
 GN RAD27 OR RTH1 OR YKL113C OR YKL510.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-92221689; PubMed=1561835;
RA Jacquier A., Legrain P., Dujon B.;
RT "Sequence of a 10.7 kb segment of yeast chromosome XI identifies the
RT APN1 and the BAF1 loci and reveals one tRNA gene and several new open
RT reading frames including homologs to RAD2 and kinases.";
RL Yeast 8:121-132(1992).
RN [2]
RP CHARACTERIZATION.
RX STRAIN-BALB/C;
RC MEDLINE-95011546; PubMed=7926735;
RA Harrington J.J., Lieber M.R.;
RT "Functional domains within FEN-1 and RAD2 define a family of
RT structure-specific endonucleases: implications for nucleotide
RT excision repair.";
RL Genes Dev. 8:1344-1355(1994).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE-95113773; PubMed=7814325;
RA Reagan M.S., Pittenger C., Siede W., Friedberg E.C.;
RT "Characterization of a mutant strain of *Saccharomyces cerevisiae* with
RT a deletion of the RAD27 gene, a structural homolog of the RAD2
RT nucleotide excision repair gene.";
RL J. Bacteriol. 177:364-371(1995).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S93804; AAB21998.1; --
CC EMBL: Z28113; CAAB1953.1; --
CC PIR: S22267; S22267.
CC HSP: Q58839; 1A76.
CC SGD: S0001596; RAD27.
CC InterPro: IPR002421; 5_3_exonuclease.
CC InterPro: IPR005813; Exo-N.I.
CC InterPro: IPR003584; HHH-2.
CC InterPro: IPR001532; XPGC_Rad.
CC Pfam: PF00752; XPG_N; 1.
CC Pfam: PF00867; XPG_I; 1.
CC Pfam: PF01367; 5_3_exonuclease; 1.
CC PRINTS: PR00853; XPGRADSUPER.
CC SMART: SM00279; RHH2; 1.
CC SMART: SM00484; XPGI; 1.
CC SMART: SM00485; XPGN; 1.
CC PROSITE: PS00841; XPG_1; 1.
CC PROSITE: PS00842; XPG_2; 1.
KW DNA repair; Hydrolase; Nuclease; Nuclear protein.
FT DOMAIN 1 105 N-DOMAIN.
FT DOMAIN 120 251 I-DOMAIN.
SQ SEQUENCE 382 AA; 43279 MW; 1F54B08720121C8C CRC64;

Query Match 48.4%; Score 938; DB 1; Length 382;
Best Local Similarity 49.5%; Pred. No. 3e-54;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

OY 1 MGKGLTKLIADNAPKAKKEKFESYGRKIAVDASMSIVQFLIVVGRIGMETLTNEAGE 60
DB 1 MGKGLNATISEHVPISATRKSDIKSFGRKVAIDASMSIVQFLIVVGRIGMETLTNEAGE 60
OY 61 VTSLGGMNRTIRLEAGIKPVVFDGKPPDMKKOELAKRYSKRDATKDLTAVEVGD 120
DB 61 TTSLGGMNRYRLMDINGIKPICYFDGKPPDLKSHLFRSRRVVEKKLAETTELE 120
OY 121 KDAIEKLSKRTVKTVOHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 180

Db 121 K--MKOERRLVKSKHEHNEAQKLLGLMGIPYIAPTEAEQAQCAELAKKGVVAAASD 177
OY 181 MDSITFCAPFRLHLMDSKKIPVMEFDVAKVLELELTMDQFDLCLTLCGLGVCDSIK 240
DB 178 MDVLCYRTFPLRLHLSFSEAKKEPIHEIDTFLVRLGLDLTIEQFVLDLMLGCDYCESIR 237
OY 241 GIGGQALKLIRHOGSIESILENL-----NKDRYQIPEDPYPQEARLKEPNVTLDIPE 295
DB 238 GVGPTVALKLIKTHSGISIEKIVFIESGESNNTRKWKIPEDPYPQEARLKEPNVTLDIPEV-IGNE 296
OY 296 --LKVTPADEGLISFLVKDNGFNEDRVTKAIEKIKSAKNSQGRLESFKEPTATT--- 350
DB 297 INLKWSPPKKEELIYLCDDKFKSEERVKSGISRLKGLKGLTQGRDLGDFQVVPKTEQ 356
OY 351 -SAPLKRKETSDKTSKAAANKTK 373
DB 357 LAAAKRAQENKKLNK-NKNKVTK 379

RESULT 5
FEN_MENJA STANDARD; PRT; 326 AA.
ID FEN_MENJA
AC Q58839;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flap endonuclease.
GN FEN OR MJ1444.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed=6689087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.W., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
RT *jannaschii*.";
RL Science 273:1058-1073(1996).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE-98440432; PubMed=9765572;
RA Rao H.G., Rosenfeld A., Wetmur J.G.;
RT "Methanococcus jannaschii flap endonuclease: expression, purification,
RT and substrate requirements.";
RL J. Bacteriol. 180:5406-5412(1998).
RN [3]
RP CHARACTERIZATION, AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE-99400693; PubMed=10102570;
RA Bae K.W., Bae K.W., Cho C.S., Hwang K.Y., Kim H.-R., Sung H.-C.,
RA Cho Y.;
RT "Expression, purification, characterization and crystallization of
RT flap endonuclease-1 from *Methanococcus jannaschii*.";
RL Mol. Cells 9:45-48(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE-98363215; PubMed=9699635;
RA Hwang K.Y., Bae K., Kim H.-Y., Cho Y.;
RT "The crystal structure of flap endonuclease-1 from *Methanococcus*
RT *jannaschii*.";
RL Nat. Struct. Biol. 5:707-713(1998).
CC -1- FUNCTION: Has 5' endor/exonuclease and 5' pseudo-Y-endonuclease
CC activities. Cleaves the junction between single and double-
CC stranded regions of flap DNA.

```

CC CC      -!- COFACTOR: Binds 2 magnesium ions per subunit.
CC CC      -!- MISCELLANEOUS: ITS OPTIMUM PH IS 6-7.
CC CC      -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC CC      SUBFAMILY.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL: U67585; AAB9454.1; .
CC CC      PDB: 1A76; 03-AUG-99.
CC CC      PDB: 1A77; 03-AUG-99.
CC CC      TIGR: MJ144.
CC CC      InterPro: IPR000513; Exo_N.I.
CC CC      InterPro: IPR001532; XPGC_Rad.
CC CC      Pfam: PF00752; XPG_N.1.
CC CC      Pfam: PF00867; XPG_I.1.
CC CC      PRINTS: PR00853; XPGRADSUPER.
CC CC      SMART: SM00279; HH2.1.
CC CC      SMART: SM00484; XPGI.1.
CC CC      SMART: SM00485; XPGN.1.
CC CC      Hydrolase; Nuclease; Endonuclease; Magnesium; Metal-binding.
CC CC      3D-structure; Complete proteome.
CC CC      METAL 154 154 MAGNESIUM 1.
CC CC      SEQUENCE 326 AA; 37527 MW; 311427F2B4B67580 CRC64;
CC CC      Query Match 28.9%; Score 560.5; DB 1; Length 326;
CC CC      Best Local Similarity 39.1%; Pred. No. 9.5e-30;
CC CC      Matches 127; Conservative 68; Mismatches 114; Indels 16; Gaps 7;
CC CC      QY 23 FESVFGKIAVDASMSIYQFLIVGRTGMEVTLTNEAGEVTSHQGMFETILLLEAGIKP 82
CC CC      Db 16 FEDLKKKVALDGMNLYQFLSIRLDSPLRNKKEIYSANVGYFTIHLLENDITP 75
CC CC      QY 83 VYVFDGKPPDMKQELAKRYKREDATKDLTEAVEVGDKDAIKELSKRVKVTROHNEDC 142
CC CC      Db 76 IIVFDGEPKPKATKTRVREKREKELKREKKEKEDFEAAKYAKRVSYLTPKRWENC 135
CC CC      QY 143 KELLRLMGVPVPAEPAEAEACALCINDKVFVAVASDMSLTFGAPRFLRLMDPSSK 202
CC CC      Db 136 KYLLSLMGIPVPAEPEGEQAQASYWAKKGDYVAVVSYQDYLALYGAPRVYRNL--TTTKE 193
CC CC      QY 203 IPVMEFDVAKVLELTMDQFIDLCILCGDY-CDSIKIGGQTKALKIRHQHSIESIL 261
CC CC      Db 194 MPFL-ELNEVLEDLRLISLDDLLIDIAIFMGTDYNPNGVKGIGPKRAYELVRS-GVAKDVL 251
CC CC      QY 262 ENLNKRDYQIPEDWPYQEARLFEPNVTLDIPELKWTPADEGLISFLVKONGFNEDRV 321
CC CC      Db 252 KK-----EVEYDEIKRIKPEPKVT-DNYSLSLKLDPKKEGIKFLVDENFDYDRV 301
CC CC      QY 322 TKALEIKTS-AKNKSQGRLESFEK 345
CC CC      Db 302 KKHVDKLYLNANKTKQTLDAWFK 326
CC CC      RESULT 6
CC CC      EXOL_SCHPO STANDARD: PRT; 571 AA.
CC AC      P53695;
CC DT      01-OCT-1996 (Rel. 34, Created)
CC DT      01-OCT-1996 (Rel. 34, Last sequence update)
CC DT      15-JUN-2002 (Rel. 41, Last annotation update)
CC DE      Exonuclease I (EXO I).
CC GN      EXO1 OR SPBC29A10.05.
CC OS      Schizosaccharomyces pombe (Fission yeast).
CC OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
CC OC      Schizosaccharomycetes.

```

```

OX CC      NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 139-160.
RX STRAIN=972;
RX MEDLINE=95158876; PubMed=7855597;
RT "A role for exonuclease I from S. pombe in mutation avoidance and
RT mismatch correction."
RL Science 267:1166-1169(1995).
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat M., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharf S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochard M., Gallard C., Tallada V.A., Carzon A., Thode G.,
RA Degu R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carruttil L., Lowe J., McCombie W.K., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT The genome sequence of Schizosaccharomyces pombe.
RL Nature 415:871-880(2002).
CC CC      -!- FUNCTION: 5'->3' DOUBLE-STRANDED DNA EXONUCLEASE THAT COULD ACT IN
CC CC      -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC CC      -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL: L35174; AAC1648.1; .
CC CC      EXBL: AL034463; CA22433.1; .
CC CC      InterPro: IPR000513; EXO_N.I.
CC CC      InterPro: IPR003584; HH2.1.
CC CC      InterPro: IPR001440; TPR.
CC CC      InterPro: IPR001532; XPGC_Rad.
CC CC      Pfam: PF00752; XPG_N.1.
CC CC      Pfam: PF00867; XPG_I.1.
CC CC      PRINTS: PR00853; XPGRADSUPER.
CC CC      SMART: SM00279; HH2.1.
CC CC      SMART: SM00484; XPGI.1.
CC CC      SMART: SM00485; XPGN.1.
CC CC      PROSITE: PS00841; XPG.1.1.
CC CC      PROSITE: PS00842; XPG.2.1.
CC CC      DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
CC CC      DOMAIN 1 96 N-DOMAIN.
CC CC      FT DOMAIN 114 245 I-DOMAIN.
CC CC      SEQUENCE 571 AA; 63867 MW; 7D21B3778104BA9D CRC64;
CC CC      Query Match 13.4%; Score 260.5; DB 1; Length 571;

```

```

Best Local Similarity 25.4%; Pred. No. 6.6e-10;
Matches 110; Conservative 73; Mismatches 167; Indels 83; Gaps 17;

QY 1 MGKIGTKLADNAPAM-KEOKFESYFGKIAVDASMSIYQFLVVGRTGMEPLTNEAG 59
DB 1 MGKIGLLGLL-----KPMQKSHVVEESGKTLGVGVVWLHKAVITCAH-----ELAFNK-- 50
QY 60 EVTSHLOGFMRTIRLLRAGLKPYVVDGKPPDMKKQELAKYSKRDATKDLTAEVVG 119
DB 51 ETKYLKYLHQAHLQYGVPLVDFDGGPLCKASTQKKRRQ-----EAFELG 103
QY 120 DK-----DAIEKLSKRTVKVTRQHNEDCKRLRLMVGVPVVEAPSBAEACALCIND 171
DB 104 KKLWDEKKSQAIMQFS-KCDVTPEMANKLILALREHIESIVAPYEAQAQLVLEKEN 162
QY 172 KVFVASEDMDSITGAPRFLRHLMDPSSKKIPVMEFDVAKYLE-ELBTLMDQFDICIL 230
DB 163 IIDGITEDSDMLVEGAQTVL-FKMDGFGNCITIRNDIANAOQDLNLRPIEKLURHMAIF 221
QY 231 CGGDCYCDISKIGGOTALKLRQHGSIESTILEMLNKDR-YQIPEDWYOEAE-----R 281
DB 222 SGCDYTDGVAGMGLKTLRYLQKYPEPRAAIRAMRLDKSLKVPVSFEKEFALADLAFRHQ 281
QY 282 RLKPEPNVT---LDIPELKWTPADEGLISFLVKONGF-----NEDRVTKATEIKTSA 331
DB 282 RVYCPKDKTLVHLSPEBELSVHEDAFIGSFF-DNQLAIDIAEGRSNPIKCAFDIKDS 339
QY 332 KNKS-----SQGRLESFF-----KPTATTS-----APLKRKE 358
DB 340 SMOFSQTKTITTSKRGISKTDISNFFMKSIPTSPSKRPTKSTSLIDVTWVKVQRTHLANDI 399
QY 359 TSKDTSKAAANK 371
DB 400 SSKQSIKSANEK 412

RESULT 7
ID YA31_SCHPO STANDARD; PRT; 496 AA.
AC Q05708:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C18B11.01c in chromosome 1.
GN SPAC18B11.01c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spours J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris B., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Coiffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

```

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shekovskii G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z50728; CAA90586.1; -
DR InterPro; IPR000513; Exo_N1.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF00752; XPG_N; 1.
DR Pfam; PF00867; XPG_I; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00484; XPG1; 1.
DR SMART; SM00485; XPGN; 1.
DR Hypothetical protein; Hydrolase; Nuclease; Endonuclease.
RW SEQUENCE 496 AA; 57763 MW; FFA16D26846F24C3 CRC64;
SQ
Query Match 12.7%; Score 245.5; DB 1; Length 496;
Best Local Similarity 30.2%; Pred. No. 5.3e-09;
Matches 62; Conservative 47; Mismatches 77; Indels 19; Gaps 5;
QY 111 DLTEAVEVGDKDAIEKLSKRTVKVTRQHNEDCKRLRLMVGVPVVEAP--SBAEACALC 168
DB 251 ELQKAIKLKQTE-LDKLERLYRPSQNIFFELIKLGPASFPIGVEAEAFASIS 309
QY 169 INDKVFAVASEDMDSITGAPRFLRHLMDPSSKKIPVMEFDVAKYLEELTMDQFDIC 228
DB 310 QNNLAYAVATQTDVLLGSSMISNFDLNDNFHLPLQINDPKIAQELNLTDFGQDYC 369
QY 229 ILCCGDCYCDISKIGGOTALKLRQHGSIESTILEMLNKDR-YQIPEDWYOEAEARL 285
DB 370 LMCQGDFTSRIPKIGPVRLKIRYGNADFVLRKALNVEEKYIIPDYIKKFLTKKLFT 429
QY 286 EPNVTLDIPELKWTPADEGLISFL 310
DB 430 D-----LPSNNELFSFI 441
RESULT 8
RAD2_YEAST
ID RAD2_YEAST STANDARD; PRT; 1031 AA.
AC P07276;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA repair protein RAD2.
GN RAD2 OR YGR256C.
OS Saccharomyces cerevisiae (baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86083164; PubMed=3000874;
RA Nicolet C.M., Chenevert J.M., Friedberg E.C.;
RA "The RAD2 gene of Saccharomyces cerevisiae: nucleotide sequence and
RA transcript mapping.";
RL Gene 36:225-234(1985).
RN [2]

```


RA Clarkon S.G.;
 RT "The human gene for xeroderma pigmentosum complementation group G
 RL (XPG) maps to 13q33 by fluorescence in situ hybridization.";
 RN Genomics 21:283-285(1994).
 [5]
 RP CHARACTERIZATION.
 RX MEDLINE=94266772; PubMed=8206890;
 RA O'Donovan A., Scherly D., Clarkson S.G., Wood R.D.;
 RT "Isolation of active recombinant XPG protein, a human DNA repair
 RL endonuclease.";
 RN J. Biol. Chem. 269:15965-15968(1994).
 [6]
 RP CHARACTERIZATION.
 RX MEDLINE=94376893; PubMed=8090225;
 RA O'Donovan A., Davies A.A., Moggs J.G., West S.C., Wood R.D.;
 RT "XPG endonuclease makes the 3' incision in human DNA nucleotide
 RL excision repair.";
 RN Nature 371:432-435(1994).
 [7]
 RP CHARACTERIZATION.
 RX MEDLINE=94359802; PubMed=8078765;
 RA Habraken Y., Sung P., Prakash L., Prakash S.;
 RT "Human xeroderma pigmentosum group G gene encodes a DNA
 RL endonuclease.";
 RN Nucleic Acids Res. 22:3312-3316(1994).
 [8]
 RP REVIEW ON VARIANTS XP-G.
 RX MEDLINE=93374920; PubMed=10447254;
 RA Cleaver J.E., Thompson L.H., Richardson A.S., States J.C.;
 RT "A summary of mutations in the UV-sensitive disorders: Xeroderma
 RL pigmentosum, Cockayne syndrome, and trichothiodystrophy.";
 RN Hum. Mutat. 14:9-22(1999).
 [9]
 RP VARIANT XP-G VAL-792, AND VARIANT ASP-1104.
 RX MEDLINE=95038755; PubMed=7951246;
 RA Nussipikel T., Clarkson S.G.;
 RT "Mutations that disable the DNA repair gene XPG in a Xeroderma
 RL pigmentosum group G patient.";
 RN Hum. Mol. Genet. 3:963-967(1994).
 [10]
 RP VARIANT XP-G VAL-792.
 RX MEDLINE=97250499; PubMed=9096355;
 RA Nussipikel T., Lalle P., Leadon S.A., Cooper P.K., Clarkson S.G.;
 RT "A common mutational pattern in Cockayne syndrome patients from
 RL Xeroderma pigmentosum group G: implications for a second XPG
 RN function.";
 RN Proc. Natl. Acad. Sci. U.S.A. 94:3116-3121(1997).
 CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
 CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
 CC REPAIR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DISEASE: COMPLEMENTS THE DEFECT IN XERODERMA PIGMENTOSUM (XP)
 CC GROUP G (XP-G). AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY
 CC HYPERSENSITIVITY OF THE SKIN TO SUNLIGHT FOLLOWED BY HIGH
 CC INCIDENCE OF SKIN CANCER AND FREQUENT NEUROLOGIC ABNORMALITIES.
 CC -1- DISEASE: ALSO INVOLVED IN COCKAYNE'S SYNDROME (CS); AN AUTOSOMAL
 CC RECESSIVE DISEASE WHICH IS CHARACTERIZED BY A UV-SENSITIVE SKIN
 CC (WITHOUT PIGMENTATION ABNORMALITIES), NEUROLOGICAL DYSFUNCTION
 CC DUE TO DEMYELINATION OF NEURONS AND CALCIFICATION OF BASAL GANGLIA
 CC (PSYCHOMOTOR RETARDATION, DEAFNESS, OPTIC ATROPHY, RETINAL
 CC PIGMENTATION AND HYPERREFLEXES), AND DYSMORPHIC DWAFTISM (IMMATURE
 CC SEXUAL DEVELOPMENT AND MICROCEPHALY).
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
 CC SUBFAMILY.
 CC -1- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/Chromancer/Genes/XPGID300.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL: X69978; CAA49598.1;
 DR EMBL: D16305; BAA03812.1;
 DR EMBL: L20046; AAC37533.1;
 DR EMBL: X71341; CAA50481.1;
 DR EMBL: X71342; CAA50481.1; JOINED.
 DR FIR: S35993; S35993.
 DR Genew: HGNC:3437; ERCC5.
 DR MIM: 135330;
 DR MIM: 278780;
 DR InterPro: IPR000513; Exo_N_I.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR001044; XPGC_DNA_repair.
 DR InterPro: IPR001532; XPGC_Rad.
 DR Pfam: PF00752; XPG_N; 1.
 DR Pfam: PF00867; XPG_I; 1.
 DR PRINTS: PR00853; XPGRADSUPER.
 DR SMART: SM00279; Hhh2; 1.
 DR SMART: SM00484; XPGI; 1.
 DR SMART: SM00485; XPGN; 1.
 DR TIGRGRAMS: TIGR00600; rad2; 1.
 DR PROSITE: PS00841; XPG_1; 1.
 DR PROSITE: PS00842; XPG_2; 1.
 DR DNA repair; DNA-binding; Nuclear protein; Hydrolase; Nuclease;
 KW Endonuclease; Xeroderma pigmentosum; Cockayne's syndrome;
 KW Polymorphism; Disease mutation.
 FT DOMAIN 1 95 N-DOMAIN
 FT DOMAIN 753 881 I-DOMAIN
 FT DOMAIN 1057 1073 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT VARIANT 254 254 M -> V.
 FT VARIANT 792 792 /FTID-VAR_007732.
 FT VARIANT 792 792 A -> V (IN XP-G; MILD FORM).
 FT VARIANT 1104 1104 /FTID-VAR_007733.
 FT VARIANT 1104 1104 H -> D.
 FT VARIANT 1104 1104 /FTID-VAR_007734.
 FT CONFLICT 55 55 L -> P (IN REF. 2).
 FT CONFLICT 120 122 KTA -> QTS (IN REF. 2).
 FT CONFLICT 126 126 K -> Q (IN REF. 2).
 FT CONFLICT 264 266 ROY -> SSH (IN REF. 2).
 FT CONFLICT 760 760 I -> F (IN REF. 2).
 FT CONFLICT 796 796 I -> V (IN REF. 2).
 FT CONFLICT 864 872 EGPIPTGCV -> GNTNGLC (IN REF. 2).
 FT CONFLICT 959 959 R -> S (IN REF. 2).
 FT CONFLICT 959 959 R -> S (IN REF. 2).
 SQ SEQUENCE 1186 AA; 133328 MW; 8702EC68E080F1C4 CRC64;
 Query Match 11.4%; Score 220.5; DB 1; Length 1186;
 Best Local Similarity 21.3%; Pred. No. 6.6e-07;
 Matches 94; Conservative 71; Mismatches 149; Indels 127; Gaps 15;
 QY 15 PKAMKQKESYFGRKIAVDASMSIYOFLLVVGRTGMEVLTNEAGEVTSHLQGMFNFTIR 74
 DB 635 PKAVEPWEIDS-----ESESDESGSFIEV-----EQAEFPET-- 673
 QY 75 LLEAGIKPVVVDGKPPDMKQK-----LAKRYSKRDDATKDLTEAVEV 118
 DB 674 -----SKPPEQEEELVTRGEAPAESESLLDNSERDDVDGEPQEA--- 717
 QY 119 GDKDAIEKL-----SKRVKVTROHNE-----DCKRL 146
 DB 718 -EKDAESLHEWODINLELETLESNLLAQNSLAKQKQKQRIAAVTGOMFESQELL 776
 QY 147 RLGMGVFVPEAPSEAEACALCINDKVFVAVASEDMDSLTFFGAPRFLRLHMDPSSKIPVM 206
 DB 777 RLFGIPYIQAPMEAEACAIIDLTDQTSITDDSDILWLFGARHYVRNF--NKNKFE 833
 QY 207 EFDVAVKVELELTMDQIDILCGDCYDSIKGIGOTALKLIQ-----HGSESTILE- 262
 DB 834 YYQYVDFHQLGLDRNKLINLAYLLGSDYEGIPVGVCTAMEILNEFFPGH-LEPLK 892
 QY 263 -----NLNKD--RYQIPEDWPYQEARLKEPNVTLDIPELKWTA 300
 DB 893 SEWHEAQNKPIRPNPHDTPVKKKLTLOLTGFPNPAVAEAYLKPVDVDSKGSFLMGK 952

QY 301 PDEGLISPLVDKNGEDNRYTKAI-EKISAKNKSQORLESFFKPTATTISAPLKKKET 359
 Db 953 PDLDKIREFCQRYFWNTKTDESFLPVLAQDQOQLRIDSFFRLAQOKEKDAKRRIKS 1012
 QY 360 S-----DKTSKAANK 370
 Db 1013 QRLNRAVTCMLRKEKAASE 1033

RESULT 10
 XPG_MOUSE
 ID XPG_MOUSE STANDARD; PRT; 1170 AA.
 AC P35689; Q61528; Q64248;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA-repair protein complementing XP-G cells homolog (Xeroderma
 pigmentosum group G complementing protein homolog) (DNA excision
 repair protein ERCC-5).
 DE ERCC5 OR XPG OR ERCC-5.
 GN Mus musculus (Mouse).
 OS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE-94173288; PubMed-7510366;
 RA Shiomura T., Harada Y.-N., Saito T., Shiomura N., Okuno Y., Yamaizumi M.;
 RT "An ERCC5 gene with homology to yeast RAD2 is involved in group G
 Xeroderma pigmentosum.";
 RL Mutat. Res. 314:167-175(1994).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/10; TISSUE=Liver;
 RX MEDLINE-96070433; PubMed-7590748;
 RA Harada Y.N., Matsuda Y., Shiomura N., Shiomura T.;
 RT "Complementary DNA sequence and chromosomal localization of xpg, the
 mouse counterpart of human repair gene XPG/ERCC5.";
 RL Genomics 28:59-65(1995).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=DRB/2;
 RX MEDLINE-96359149; PubMed-8703115;
 RA Ludwig D.L., Mudgett J.S., Park M.S., Perez-Castro A.V.,
 RA MacInnes M.A.;
 RT "Molecular cloning and structural analysis of the functional mouse
 genomic xpg gene.";
 RL Mamm. Genome 7:644-649(1996).
 CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
 EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
 REPAIR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16306; BAA03813.1; --
 DR EMBL: U40796; AAA91039.1; --
 DR EMBL: U40795; AAB17885.1; --
 DR EMBL: U39892; AAB17885.1; JOINED.
 DR EMBL: U39893; AAB17885.1; JOINED.
 DR EMBL: U39894; AAB17885.1; JOINED.
 DR EMBL: U39896; AAB17885.1; JOINED.
 DR EMBL: U40073; AAB17885.1; JOINED.
 DR EMBL: U40431; AAB17885.1; JOINED.

DR EMBL: U40432; AAB17885.1; JOINED.
 DR EMBL: U40668; AAB17885.1; JOINED.
 DR EMBL: U40669; AAB17885.1; JOINED.
 DR EMBL: U40670; AAB17885.1; JOINED.
 DR EMBL: U40792; AAB17885.1; JOINED.
 DR EMBL: U40793; AAB17885.1; JOINED.
 DR EMBL: U40794; AAB17885.1; JOINED.
 DR MGD; MGI:103582; Erc5.
 DR InterPro; IPR000513; EXO_N_I.
 DR InterPro; IPR003584; HHH_2.
 DR InterPro; IPR01044; XPGC_DNA_repair.
 DR InterPro; IPR01532; XPGC_Rad.
 DR Pfam; PF00752; XPG_N; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00484; HH2; 1.
 DR SMART; SM00485; XPGI; 1.
 DR TIGRFAMs; TIGR00600; rad2; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 DR DNA repair; DNA-binding; Nuclear protein; Hydrolase; Nuclease;
 KW Endonuclease.
 FT DOMAIN 1 95 N-terminus
 FT DOMAIN 752 882 I-DOMAIN
 FT DOMAIN 1049 1065 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT VARIANTS 388 488 R -> C (IN STRAIN DBA/2).
 FT VARIANTS 488 488 S -> R (IN STRAIN DBA/2).
 FT VARIANTS 688 688 I -> T (IN STRAIN DBA/2).
 FT VARIANTS 1015 1015 S -> T (IN STRAIN DBA/2).
 FT VARIANTS 1021 1021 M -> I (IN STRAIN DBA/2).
 FT VARIANTS 1121 1121 S -> P (IN STRAIN DBA/2).
 FT CONFLICT 227 227 N -> M (IN REF. 1).
 FT CONFLICT 249 249 N -> M (IN REF. 1).
 FT CONFLICT 300 302 VMD -> MDE (IN REF. 1).
 FT CONFLICT 313 313 N -> M (IN REF. 1).
 FT CONFLICT 320 320 N -> M (IN REF. 1).
 FT CONFLICT 393 399 D -> DVDTGG (IN REF. 1).
 FT CONFLICT 408 408 N -> NSASEVIGPV (IN REF. 1).
 FT CONFLICT 581 581 M -> V (IN REF. 1).
 FT CONFLICT 795 795 A -> AMEKEFEL (IN REF. 1).
 FT CONFLICT 1039 1039 S -> SD (IN REF. 1).
 FT CONFLICT 1134 1134 KL -> RR (IN REF. 1).
 FT CONFLICT 1157 1158
 SQ SEQUENCE 1170 AA; 130864 MW; 4058F07FED1770ED CRC64;
 Query Watch 11.3%; Score 220; DB 1; Length 1170;
 Best Local Similarity 21.6%; Pred. No. 7e-07;
 Matches 88; Conservative 71; Mismatches 156; Indels 92; Gaps 13;
 QY 13 NAPKAMKEKESYFGKRIADVASMSYQFLIVVGTGTMETLTNEAGVTSHL----- 65
 Db 633 SAPKPMGPMES-----ESESQSGFIEVQSVVSVNSELOTESSEATHEKDAEE 684
 QY 66 -----QGMFNRTLLGAGIKPVYVDGKPPDMKKQELAKRYKRDATKDLT---EAV 116
 Db 685 PREILEEGYSRDETCILQ-----DSSDIEAMEGHEGREADDAEDMNEWQDI 730
 QY 117 EYGDQDAIEK---LSKRVKVTROHNE-----DKRLLRLMGVZVVPAPSEAE 161
 Db 731 NLEELDALESNLLARQNSLKAQKQODRIAASVTGOMFLESQELLRLFGVPTICAPMEAE 790
 QY 162 AECAALCINDKVFVAVASDMLTFGAPRFLHLMOPSSKIPVME-FDAKVLLELELT 220
 Db 791 AACAMLDLSQTSITDDSDINLFGA---RHVYKFNFNKKNKEVEYYQYVDFYSQLGLD 846
 QY 221 MOOFIDLCILGCDYCSIKIGGOTALKIRH--GSIESIL-----ENLNK--- 266
 Db 847 RNKLINLAYLGSYTEGIPYGVCTAMEILNEFPGRGLDPLKFESEKHWQAQNKVAAE 906
 QY 267 -----DRYOIPEDWPYQEARLFKEPNVTLIDIPELKWTAPDEEGLISFLVKDNG 315
 Db 907 NPYDTKVRKKLRKLQLTLPFGFPNPAVADAYLRPVYDDSRGSLFKGPKVDYKIRFXQRYFG 966

QY 316 FNEEDRVTKAIBKI-KSAKNSSQGRLESFPKPTATTSAPLKRKETS 361
 Db 967 WNRKMTDESPLVPLKHLNAHOTLRIDSFFR-----LAQEKQD 1005

RESULT 11

ID EXO_YEAST STANDARD; PRT; 702 AA.
 AC 01-FEB-1995 (Rel. 31, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Exonuclease I (EXO I) (DHS1 protein).
 GN EXO1 OR DHS1 OR YOR033C OR OR26.23.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE OF 220-702 FROM N.A.
 RX MEDLINE=94168519; PubMed=7764548;
 RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
 RT "Molecular cloning of a gene, DHS1, which complements a drug-
 RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
 RL Biosci. Biotechnol. 58:391-395(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / FY1679;
 RA de Haan M., Maarse A.C., Grivell L.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Landt O., Hiesel R., Unseld M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Tishkoff D.X., Boeger A.L., Filosi N., Galda G.M., Bertrand P.,
 RA Kane M.F., Kolodner R.D.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP FUNCTION
 RX MEDLINE=97265410; PubMed=9111347;
 RA Fiorentini P., Huang K.N., Tishkoff D.X., Kolodner R.D.,
 RA Symington L.S.;
 RT "Exonuclease I of Saccharomyces cerevisiae functions in mitotic
 RT recombination in vivo and in vitro.";
 RL Mol. Cell. Biol. 17:2764-2773(1997).
 CC -!- FUNCTION: 5'-3' double-stranded DNA exonuclease involved in
 CC mismatch repair and eventually also in mitotic recombination
 CC between direct repeats.
 CC -!- COFACTOR: Magnesium or manganese.
 CC -!- ENZYME REGULATION: Inactivated by calcium and zinc ions.
 CC -!- SUBUNIT: INTERACTS WITH MISMATCH REPAIR PROTEIN MSH2.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; S69545; AAC60570.1;
 DR EMBL; X87331; CAAG0749.1;
 DR EMBL; 274941; CAA99232.1;
 DR EMBL; U86134; AAB47428.1;
 DR PIR; JC2068; JC2068.
 DR SGD; S0005559; DHS1.
 DR InterPro; IPR000513; Exo_N_I.
 DR InterPro; IPR003584; HHH_2.

DR InterPro; IPR001532; XPGC_Rad.
 DR Pfam; PF00752; XPG_N; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; Hh2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW DNA repair; Hydrolyase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 96 N-DOMAIN.
 FT DOMAIN 114 247 I-DOMAIN.
 FT DOMAIN 501 520 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 537 553 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 618 625 ASP-RICH (ACIDIC).
 SQ SEQUENCE 702 AA; 80162 MW; 78E22F6B265DB3AA CRC64;

Query Match 11.2%; Score 218; DB 1; Length 702;

Best Local Similarity 23.7%; Pred. No. 5e-07;
 Matches 99; Conservative 70; Mismatches 176; Indels 72; Gaps 14;

QY 1 MGKIG-LTKLLADNAPKAMKEKESYFGKIAVDASMSIYO-----FLIVVGRGTG 53
 Db 1 MGIGLLPOLKPIQNPVSLRR-----YEVEVLAIDGYAWLHRAACSCAYELAMGK----- 50
 QY 54 LTNEAGEVTSHLQGFNTIIRLEAGIKPVYVDGKPPDMKKOELAKRYSKRDDATKDLT 113
 Db 51 -----PTDKYLQFFIKRPSLLTKTKRVEPYLVDFDGAIPVKSTESKRDRKRRE-NKAIA 103
 QY 114 EAV-EVGDKAIEKLSKRTVKVTRQHNEDCKRLRLMLGVVVEAPSEAEACALCINDK 172
 Db 104 ERWAGGEKKNAMDFORCVDTPEMAKCIYCKLNGIRYIVAFEADSQMVLQKNI 163
 QY 173 VFVASEDMDSLTFGAPRFLRLMDPSS-----KKIPVMEFDVAKYLELEL 219
 Db 164 VOGLISEDSLLVFGCRLLITKNDYGECEICRDNFIKLPKKPLG-----SL 212
 QY 220 TMOQFDILCIGCDYCDISIKGIGGOTALKLIRHGSIESILENLNKD-RYOIPEDW--P 276
 Db 213 TNEEITWCLSGCDYTGIPKVLITAMKLVRRPNTIERIILSIQREGKLMIPDYINE 272
 QY 277 YQARRLFEPNV-----TLDIPELKWTAPOEGLISFLVKDNGNEEDRVTKAEIK 329
 Db 273 YEAVLAFOQVFCPIRKIVSLNEIPLYLKTESKRRLYACIGFVIHRETQKQIVH 332
 QY 330 SAKN-----KSSQGRLE--SPFKPTATTSAPLKRKETS-----KTSKAAANKTK 373
 Db 333 FDDIDHHLHLKIAQGLNPDYDFHQLANREHLQLASKSNIEFGKTNNTNSEAZVK 389

RESULT 12

ID DIN7_YEAST STANDARD; PRT; 430 AA.
 AC Q12036;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA-damage inducible protein DIN7.
 GN DIN7 OR DIN3 OR YDR263C OR YD9320B.02C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OLL;
 RX MEDLINE=97234602; PubMed=9079876;
 RA Mieczkowski P.A., Fikus M., Ciesla Z.;
 RT "Characterization of a novel DNA damage-inducible gene of
 RT Saccharomyces cerevisiae, DIN7, which is a structural homolog of the
 RT RAD2 and RAD27 DNA repair genes.";
 RL Mol. Gen. Genet. 253:655-665(1997).
 RN [2]

```

RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: 5'->3' double-stranded DNA exonuclease (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- INDUCTION: By UV light, methyl methane-sulfonate (MMS) or
CC hydroxyurea (HU), and during meiosis.
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: X90707; CAA62233.1; -
CC EMBL: 270202; CAA94102.1; -
CC EMBL: 268290; CAA92581.1; -
CC SGD: S0003671; D1N7.
CC InterPro: IPR000513; Exo_N_I.
CC InterPro: IPR000584; HHH_2.
CC Pfam: PF00752; XPG_N; 1.
CC Pfam: PF00867; XPG_I; 1.
CC PRINTS: PR00853; XPGRADSUPER.
CC SMART: SM00484; XPG1; 1.
CC SMART: SM00485; XPGN; 1.
CC PROSITE: PS00841; XPG_1; 1.
CC PROSITE: PS00842; XPG_2; 1.
CC DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
FT DOMAIN 114 247 N-DOMAIN.
FT DOMAIN 114 247 I-DOMAIN.
SQ SEQUENCE 430 AA; 49034 MR; 2BC23D30832C79E9 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 430;
Best Local Similarity 24.7%; Pred. No. 2,1e-06;
Matches 68; Conservative 53; Mismatches 129; Indels 25; Gaps 6;

QY 1 MGIGKLTLLADNAPKAMKQKESYFGRKIVADSMISYQ-----FLIVGRTGMETL 54
Db 1 MGIPGLLPQ-----KRIQKQVSLKKYMYQTALDGVWHLRASCACAFELVMNKP----- 51
QY 55 TNEAGEVTSHQGMFNRTIRLEAGIKPVVYFGKPPDMKKQELAKRYKRDATKDLTE 114
Db 52 TNK-----YLOQFIKRLQLKRLKIKPIYVFGDGLFVKNHTETRRKKLENEIMAKK 105
QY 115 AVEVGKDAIEKSKRTVKVTRHNEDECKRLRLMGVFPVVEAPSEAEACALCINDKVF 174
Db 106 LWSAGNRYNMEYFKRSVDIIPEMAKIYDYCKLHSIYIVAPFEADPQWYLEKMGILIQ 165
QY 175 AVASEDMDSLTGAPRLHMDPSKKIPVMEFDVAKVLEL---ELTMDQFDLCLIC 231
Db 166 GIISEDSDLLVFGCKTLITLND-QGKALEISKDFSLPENFFLIGSELSEQGFNLVCLA 224
QY 232 GCDYCDISIKGGTALKLRHQHSIESILENLK 266
Db 225 GCDYTSIGWIKGVVTVANKIVKRYSEMKDILIQIER 259

RESULT 13
RA13_SCHPO
ID RA13_SCHPO STANDARD; PRT; 1112 AA.
AC P28706; OS9728;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA repair protein rad13.
GN RAD13 OR SPBG3E7.08C.
OS Schizosaccharomyces pombe (Fission yeast).

```

```

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC NCB1_taxid=4896;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE=93219111; PubMed=8464724;
RA Carr A.M., Sheldrick K.S., Murray J.M., Al-Harithy R., Watts F.Z.,
RA Lehmann A.R.;
RT "Evolutionary conservation of excision repair in Schizosaccharomyces
RT pombe: evidence for a family of sequences related to the
RT Saccharomyces cerevisiae RAD2 gene.";
RL Nucleic Acids Res. 21:1345-1349(1993).
RN [2]
RX SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Swilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Woesli D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lenrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedter H., Wambutt R., Purnelle B.,
RA Gorfeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.K., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: X66795; CAA47291.1; -
CC EMBL: AL023534; CAA19011.1; -
CC PIR: S22862; S22862.
CC PIR: S30301; S30301.
CC InterPro: IPR000513; Exo_N_I.
CC InterPro: IPR000584; HHH_2.
CC InterPro: IPR000593; UIM.
CC InterPro: IPR001044; XPGC_DNA_repair.
CC InterPro: IPR001532; XPGC_Rad.
CC Pfam: PF00752; XPG_N; 1.
CC Pfam: PF00867; XPG_I; 1.
CC PRINTS: PR00853; XPGRADSUPER.
CC SMART: SM00279; Hhh2; 1.
CC SMART: SM00484; XPG1; 1.

```

```

DR SMART; SM00485; XPGN; 1.
DR TIGRFAMS; TIGR00600; rad2; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR PROSITE; PS00842; XPG_2; 1.
DR DNA repair; Nuclear protein; Hydrolase; Nuclease; Endonuclease.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 742 870 I-DOMAIN.
FT CONFLICT 8 8 D -> N (IN REF. 1).
FT CONFLICT 738 743 LKNOKR -> AOKSKG (IN REF. 1).
FT CONFLICT 1112 AA; 126328 MW; 7ECF4229D5B4768 CRC64;
SQ SEQUENCE 1112 AA; 126328 MW; 7ECF4229D5B4768 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 1112;
Best Local Similarity 21.4%; Pred. No. 6.7e-06;
Matches 88; Conservative 68; Mismatches 148; Indels 107; Gaps 15;

QY 18 MKEQKESYGRKIV--DASMSITQFLIVGRTGTMETLNEAGEVTS-----63
Db 604 MKEQADDGTVTNPLNVSDDKMSVY---LLSENKDT-----GDIKSESIDAVLPTLET 655
QY 64 -----HLOGMNFRTIRLEAGIKPVYFDGKPPDKKOE---LAKRYSKRDDA 108
Db 656 SPSLSIPTDQKEASPNKGAALSKVEPEV--EKLDEEEEMIIEMAEEREYDRF 713
QY 109 TKDLTE--AVRVGDKDAIEKL-----SKRTVK-----VTRQHNEDCKRLRLMGVPVVE 155
Db 714 VSELNORHETEWNOFAFEKRLKELKQKREKRDADDEVTVQVMIKECQELLRLFLPLIV 773
QY 156 APSEAEACALCINDKVFVASEWDSITFCAPFLR-----HMDPSSKKIP 204
Db 774 APQEAACSKLELKYDGVITDSDVFLFGTRVYRNFMNQKVFELYLMDMKREFN 833
QY 205 VMEFVAVKLEELHWDQFIDLCILCGDYCDISKIGGQTALKIRO-----253
Db 834 VNQMDLIK-----LAHLGSDYTMGLSRVGVPLAETLHFEFGDTGLFFPK 879
QY 254 -----HGSIES-----ILENLK--DRYOIPEDMPYQEARLFPKPNVTLDIPELKW 298
Db 880 KWFQRLSTGHASKNDVNTPKRKINKLVKILFSEFPNPLVDLAEYLHPAVDSDQSQFQW 939
QY 299 TAPDEGLISPLVDKNGFNDRVTRKAIEKISAKN-----SSQGRLESFFK 345
Db 940 GIPDLDELPLMATVGNKQRTNEVLLFVQDMHKKOFVGTQSNLTQFFE 990

RESULT 14
XPG_XENIA
ID XPG_XENIA STANDARD; PRT; 1196 AA.
AC P14629;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-G cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog).
GN ERCC5 OR XPG OR XPGC.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93247645; PubMed=8483504;
RA Scherly D., Nussipikel T., Corlet J., Ucla C., Bairoch A.,
RA Clarkson S.G.;
RT "Complementation of the DNA repair defect in Xeroderma pigmentosum
RL Nature 363:182-185(1993).
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
CC REPAIR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.

```

```

CC -----THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69977; CAA49597.1; -.
DR PIR; S35994; S35994.
DR InterPro; IPR000513; EXO_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001044; XPGC_DNA_repair.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PF00867; XPG_1; 1.
DR PRINTS; PF00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
DR TIGRFAMS; TIGR00600; rad2; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR PROSITE; PS00842; XPG_2; 1.
DR DNA repair; DNA-binding; Nuclear protein; Hydrolase; Nuclease;
DR Endonuclease.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 786 914 I-DOMAIN.
FT DOMAIN 936 952 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 1079 1095 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 1196 AA; 134206 MW; 1F1CE1891A3C0623 CRC64;

Query Match 10.3%; Score 200; DB 1; Length 1196;
Best Local Similarity 21.9%; Pred. No. 1.4e-05;
Matches 72; Conservative 64; Mismatches 137; Indels 56; Gaps 11;

QY 88 GPPDMKKQELAKRYSKDDATKDLTEAVEVGDKDAIEK---LSKRTVYVTRHNE----140
Db 740 GEPHEGSNEESK---DIDDVSNWQD--ISVELESLENNLYVQQTSLQAOQQQQRIRAA 795
QY 141 -----DCKRLRLMGVPVVEAPSEAEACALCINDKVFVASEWDSITFCAPFLR 192
Db 796 TVTQGMCLSEQLLQFLGIPYIVAPMEAEAOACAILDLTDQTSGITDSDIWLFGARHYV 855
QY 193 RHLMDPSSKKIPVMEFVAVKLEELHWDQFIDLCILCGDYCDISKIGGQTALKIR 252
Db 856 KNEF---SONKHVEYVQYADIHNGQLDRSKLINLAYLIGSDYTEGPTGVYVSAMEILN 912
QY 253 QH--GSIESILE-----NLNKRYQIP-----EDMPYQEARLFPK 287
Db 913 EFTGQGLEPLVFKFEWSEAKQKKMRPNPNDTKVKKLLDLDDLOQSFNPNAVASYLKP 972
QY 288 NYTLDIPELKWTPADPEGLISFLVKNGF---NEORVTKAIEKISAKNKSOGRLSEFF 344
Db 973 VVDESKSAFSGWRPDLEIQREFCESRFGYRKLKTDVLLPV--LKQLNAOQTQLRIDSDFF 1030
QY 345 KPTATTSAPLKKRKETSDDTKSAAANKTK 373
Db 1031 RLEQHEAAGLK----SQRLLRAVTCMKRK 1055

RESULT 15
DPOL_THECA
ID DPOL_THECA STANDARD; PRT; 834 AA.
AC P80194;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (TAC polymerase 1).
GN POLA.
OS Thermus caldophilus.
OC Bacteria; Thermus; Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.

```

```
OX NCBI_TaxID=272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CK24;
RA Kwon S.-T., Kim J.S., Park J.H., Kim H., Lee D.-S.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-21.
RC STRAIN=CK24;
RX MEDLINE=93285135; PubMed=8508785;
RA Park J.H., Kim J.S., Kwon S.-T., Lee D.-S.;
RT "Purification and characterization of Thermus caldophilus GK24 DNA
  polymerase.";
RL Eur. J. Biochem. 214:135-140(1993).
CC -!- FUNCTION: THERMOSTABLE ENZYME THAT HAS 5'-TO-3' EXONUCLEASE
  ACTIVITY AND NO 3'-TO-5' EXONUCLEASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
  + (DNA)(N).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U62584; AAB81398.1; -.
DR PIR; S33287; S33287.
DR HSP; P19821; 1TAO.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR001098; DNA_pol.
DR InterPro; IPR002398; DNA_pol.
DR InterPro; IPR003133; Exo_N.1.
DR InterPro; IPR003583; HHH_1.
DR InterPro; IPR003584; HHH_2.
DR Pfam; PF00476; DNA_pol_A; 1.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF02739; 5_3_exonuc_N; 1.
DR PRINTS; PR00868; DNAPOLI.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00278; Hhh1; 1.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00482; POLAC; 1.
DR TIGRfams; TIGR00593; pola; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.
FT DOMAIN 412 834 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 834 AA; 93798 MW; A851FF3C3076348E CRC64;

Query Match          9.5%; Score 184.5; DB 1; Length 834;
Best Local Similarity 25.4%; Pred. No. 9.4e-05;
Matches 78; Conservative 47; Mismatches 115; Indels 67; Gaps 13;

QY 28 GRKIAVDASMSIQFLIVVGRGTMETLNEAGEVTSHLQGMFNRITRL--EAGIKPVY-V 85
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 12 GRVLLVDGHLAYRTTF-----ALKGLTTSRGEPVQAVGYGFAKSLKALKEDGYKAVFV 66

QY 86 FDGKPPDPWKQELAKRYSKRDDATKDLTAEVGVGDKDAIEKLSKRTVVKVTRQHNEDCKRL 145
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 67 FDAKAPSRHEAY-----EAYKAGRAPTFDPFQQLALI-----REL 103

QY 146 LRLMGVPVVEAPS-EAEACALCINDKV-----FAVASEDMDSITFGAPFLR---HL 195
   : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 104 VDLGLGFRLEVPVGEADVLATLANKPEGEYEVRIITADRLDQLVSDRVAVLHPEGHL 163

QY 196 MDPSSKKIPVMEFDVARVLEELATMOQFDICILCG--CDYCDSIKIGIGGTALKLIRQ 253
   : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
DB 164 ITPE-----WLMQKYLKPEQWDFRALVGPDPNDLPCVKGICETALKLKE 211

QY 254 HGSITSLINLNKDRYQIPED-----WPYQEARLRFKE-----PNVTLIDIPELKWTAPDE 303
```

```
Db 212 WGSLENLKRLDRVK---PENVRKIKAHLEDLRSLSELSRVRTDLPLEVDLAQGREPDR 268
QY 304 EGLISFL 310
Db 269 EGLRAFL 275
```

Search completed: May 8, 2003, 21:41:34
Job time : 38 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 21:34:28 ; Search time 58 Seconds
(without alignments)
1346.411 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGIGSLTKLADNAPKAMKE.....SDRTSKAAANKTKAGGKKK 379

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1807.5	93.2	380	10 Q9SXQ6	Q9sxq6 oryza sativ
2	1313.5	67.7	362	10 Q65231	Q65231 arabidopsis
3	1069	55.1	382	13 P70040	P70040 xenopus lae
4	1068	55.1	382	13 O57351	O57351 xenopus lae
5	1068	55.1	382	13 P70054	P70054 xenopus lae
6	1063.5	54.8	381	13 Q90YB0	Q90YB0 gallus gall
7	1043	53.8	380	11 Q91Z50	Q91Z50 mus musculus
8	1042	53.7	380	11 Q8R069	Q8R069 mus musculus
9	1040	53.6	380	11 Q9JHW7	Q9JHW7 rattus norv
10	1036.5	53.5	382	5 Q9N3T2	Q9n3t2 caenorhabdi
11	1009	52.0	378	11 Q99M86	Q99M86 mus musculus
12	998.5	51.5	385	5 Q96830	Q96830 drosophila
13	887	45.7	650	5 Q9GZ01	Q9gz01 plasmodium
14	887	45.7	672	5 Q9U0K1	Q9u0k1 plasmodium
15	841.5	43.4	345	5 Q8SS91	Q8ss91 encephalito
16	648.5	33.4	340	17 Q93634	Q93634 pyrococcus

17	635	32.8	343	17 Q9V0P9	Q9v0p9 pyrococcus
18	632.5	32.6	343	17 Q50123	Q50123 pyrococcus
19	628	32.4	346	17 Q8ZIN2	Q8zyn2 pyrobaculum
20	622	32.1	338	17 Q8TXU4	Q8txu4 methanopyru
21	586.5	30.2	338	17 Q8TIY5	Q8tiy5 methanosarc
22	542	28.0	328	17 Q27670	Q27670 methanobact
23	538.5	27.8	336	17 Q29975	Q29975 archaeoglob
24	519	26.8	304	17 Q976H6	Q976h6 sulfolobus
25	514.5	26.5	401	17 Q9VEY5	Q9vfy5 aeropyrum p
26	506	26.1	302	17 Q980U8	Q980u8 sulfolobus
27	466	24.1	336	17 Q9HJDA	Q9hjd4 thermoplasma
28	446	23.0	335	17 Q97B98	Q97b98 thermoplasma
29	369	19.0	327	17 Q9HQ27	Q9hq27 halobacteri
30	336	17.3	1516	5 Q96154	Q95154 plasmodium
31	278	14.3	415	12 Q91FF5	Q91ff5 chilo iride
32	268	13.8	366	5 Q8SQX4	Q8sqx4 encephalito
33	267.5	13.8	317	10 Q9C7N8	Q9c7n8 arabidopsis
34	257	13.3	726	5 Q9U9Q6	Q9u9q6 drosophila
35	257	13.3	726	5 Q9VRJ0	Q9vrj0 drosophila
36	244.5	12.6	1037	3 Q8WZX1	Q8wzx1 neurospora
37	243.5	12.6	1257	5 Q917N6	Q917n6 drosophila
38	243	12.5	298	12 Q8QUT3	Q8qut3 infectious
39	241	12.4	298	12 Q9QTF3	Q9qtf3 red sea bre
40	240	12.4	734	13 Q9W6K2	Q9w6k2 xenopus lae
41	235	12.1	497	10 Q9LPD2	Q9lpd2 arabidopsis
42	230	11.9	837	11 Q9QZ11	Q9qz11 mus musculu
43	230	11.9	837	11 Q9Z3A5	Q9z3a5 mus musculu
44	229	11.8	390	12 Q8QY21	Q8qy21 rana tigrin
45	227	11.7	800	4 Q75466	Q75466 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9SXQ6	PRELIMINARY;	PRT;	380 AA.
AC	Q9SXQ6;			
DT	01-MAY-2000 (TReMBLrel. 13, Created)			
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)			
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)			
DE	FEF-1.			
GN	OSFEN-1.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
CC	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=20256470; PubMed=10798612;			
RA	Kimura S., Ueda T., Hatanaka M., Takenouchi M., Hashimoto J.,			
RA	Sakaguchi K.;			
RT	"plant homologue of flap endonuclease-1: molecular cloning,			
RT	characterization, and evidence of expression in meristematic			
RT	tissues.";			
RL	Plant Mol. Biol. 42:415-427(2000)			
DR	EMBL; AB021666; BAA36171.1; -			
DR	HSP; Q58839; IA76.			
DR	InterPro; IPR002421; 5_3_exonuclease.			
DR	InterPro; IPR000513; Exo_N_I.			
DR	InterPro; IPR003584; HHH_2.			
DR	InterPro; IPR001532; XPGC_Rad.			
DR	Pfam; PF01367; 5_3_exonuclease; 1.			
DR	Pfam; PF00867; XPG_I; 1.			
DR	Pfam; PF00752; XPG_N; 1.			
DR	PRINTS; PR00853; XPGRADSUPER.			
DR	SMART; SM00475; 53EXOC; 1.			
DR	SMART; SM00279; HhH2; 1.			
DR	SMART; SM00484; XPGI; 1.			
DR	SMART; SM00485; XPGN; 1.			
DR	PROSITE; PS00841; XPG_1; 1.			
DR	SEQUENCE 380 AA; 42792 MW; E0148AAFA95A7364 CRC64;			

```
Query Match          93.2%; Score 1807.5; DB 10; Length 380;
Best Local Similarity 92.1%; Pred. No. 9.8e-120;
Matches 350; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MGKIGLTKLLADNAPKAMKEQKFESYFGRIKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
DB 1 MGKIGLTKLLADNAPKAMKEQKFESYFGRIKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60

QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVDFGKPPDMKKQELAKRYSKRDDATKDLTAEVGD 120
DB 61 VTSHLQGMFNRTIRLLEAGIKPVYVDFGKPPDLKKQELAKRYSKREDATKELTAEVGD 120

QY 121 KDAIEKLSKRTYKVTQHNEDCKRLRLRMGVVPEAPSEAEACALCINDKVFVASED 180
DB 121 KDAIEKSPKRTYKVTQHNEDCKRLRLRMGVVPEAPSEAEACALCINDMVFVASED 180

QY 181 MDSLTCGAPFLRHLMDPSSKKIPVMEFDVAKVLELELTMDQFIDLCLGCGDYCDSIK 240
DB 181 MDSLTCGAPFLRHLMDPSSKKIPVMEFEVAKVLELELTMDQFIDLCLGCGDYCDSIK 240

QY 241 GIGGQATALKLIRHOGSIESILENLKRDYQIPEDWPYQEARLKEPNVTLIDIPELKWA 300
DB 241 GIGGQATALKLIRHOGSIESILENLKRDYQIPEDWPYQEARLKEPNVTLIDIPELKWA 300

QY 301 PDEGLISFLVKGNGFNEDRVTKAIEKIKSAKNKSSQGRLESFFKPTATTSAPLKRKETS 360
DB 301 PDEGLVEFLVKGNGFNEDRVTKAIEKIKFAKNKSSQGRLESFFKPVVYSTVPLKRKETS 360

QY 361 DKTSKAAANKTK-AGGKKK 379
DB 361 EKPTKAVANKTKGAGGKKK 380

RESULT 2
065251 ID 065251 PRELIMINARY; PRT; 362 AA.
AC 065251;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE F21E10.3 protein.
GN F21E10.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Davidson S., Rohlfing T., David M., O'Brian D.;
RT "The sequence of A. thaliana F21E10."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Washu;
RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wilson R.;
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF058914; AAC13596.1;
RL HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR
```

```
DR InterPro; IPR000513; Exo_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR SEQUENCE 362 AA; 41205 MW; A4DD1706C3AD2D9D CRC64;

Query Match          67.7%; Score 1313.5; DB 10; Length 362;
Best Local Similarity 72.1%; Pred. No. 6.9e-85;
Matches 259; Conservative 30; Mismatches 33; Indels 37; Gaps 3;

QY 1 MGKIGLTKLLADNAPKAMKEQKFESYFGRIKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
DB 1 MGKIGLTKLLADNAPKAMKEQKFESYFGRIKIAVDASMSIYQFLIVRAPSSPL-LLEELGL 59

QY 61 VTS-----HLOGMFNRTIRLLEAGIKPVYVF 86
DB 60 KCSLMKLVSLWIFLDCVLCIYVFFVWVYMMFNRLHOGMFNRTIRLLEAGIKPVYVF 119

QY 87 DGKPPDMKKQELAKRYSKRDDATKDLTAEVGVGDKAIEKLSKRTYKVTQHNEDCKRL 146
DB 120 DGKPPELRQELAKRYSKRADATADLTGAEAGNKEDIEKYSKRTYKVTQHNEDCKRL 179

QY 147 RLMGVVPEAPSEAEACALCINDKVFVASEDMDSLTFGAPFLRHLMDPSSKKIPVM 206
DB 180 RLMGVVPEATSEAEACALCCKGKVGAVASEDMDSLTFGAPFLRHLMDPSSKKIPVM 239

QY 207 EFDVAKVLELELTMDQFIDLCLGCGDYCDSIRGIGQTALKLIRHOGSIESILENLK 266
DB 240 EFEVAKILELTMDQFIDLCLGCGDYCDSIRGIGQTALKLIRHOGSIESILENLK 299

QY 267 DRYQIPEDWPYQEARLKEPNVTLIDIPEL--KWTAPDEEGLISFLVKGNGFNEDRVTK 323
DB 300 ERYQIPPEWPYNEARKLFKEPVDITDESQDKTWTSPDEGIVQFLVNGFNIDRVTK 358

RESULT 3
P70040 ID P70040 PRELIMINARY; PRT; 382 AA.
AC P70040;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE 5' nuclease XFENIA.
GN FEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Piploidea;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC Bibikova M., Chi E., Wu B., Kim K., Carroll D.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Li J.-L., Cox L.S.;
RT "Cloning and investigation of Xenopus Fen1: developmental expression
and function in DNA replication."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64563; AAB06176.1;
DR EMBL; AF065397; AAD02814.1;
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
```

DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRAUSUPER.
 DR SMART; SM00279; HH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR Endonuclease.
 KW Endonuclease.
 SQ SEQUENCE 382 AA; 42668 MW; 9B1DB0EDAD158D57 CRC64;

Query Match 55.1%; Score 1069; DB 13; Length 382;
 Best Local Similarity 55.1%; Pred. No. 1.4e-67;
 Matches 210; Conservative 63; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGKGLTKLLADNAPKAMKEQKFSYFGKRIYAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
 DB 1 MGHLGLAKLIADVAPAAIKHEHDIKSYFGKRVAVDASMSIYQFLIVVGRGTGNETLTNEAGE 59
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
 DB 60 TTSHLGMGFYRTIRVHEGIRKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 119
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPVVEAPSEAEAEACALCINDKVFVASED 180
 DB 120 VENIEKTRKLVKVTQHNEDCKRLRLMGVPVVEAPSEAEAEACALCINDKVFVASED 179
 QY 181 MSLIFGAPRFLRLHMDSPSKIPVMEFDVAKVLELELTMDQFIDLCILCGCDYCDSTK 240
 DB 180 MDALTFTGTVLLRHHTASEAKKLPIQEFHLNRVTDIGITHEQFVDFDLCILGSDYCTIR 239
 QY 241 GIGGQATKLIROHGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVT-LDPELKW 299
 DB 240 GIGPKRAIDLROHKTIDEITIDNLDKYPVPENWLHKEAKHLFLEPEVVDITELKWI 299
 QY 300 APDEGLVAFNGEKFQFSEDRIRNGAKLAKNRQSTQGRLDFFKVTGSVSS-TKRKEA 358
 DB 300 EPDEGLVAFNGEKFQFSEDRIRNGAKLAKNRQSTQGRLDFFKVTGSVSS-TKRKEA 358
 QY 360 SDKTSKAAANKTKAGG 376
 DB 359 E---SKGSAKKAKTGG 372

RESULT 4
 O57351 PRELIMINARY; PRT; 382 AA.
 AC O57351
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Flap endonuclease 1.
 GN FEN1.

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98204872; PubMed=9535864;
 RA Kim K., Biade S., Matsumoto Y.,
 RT "Involvement of flap endonuclease 1 in base excision DNA repair."
 RL J. Biol. Chem. 273:8842-8848(1998).
 DR EMBL; AF036327; AAB88707.1;
 DR HSSP; Q58839; 1A76
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N.1.
 DR InterPro; IPR003584; HHH_2.
 DR InterPro; IPR001532; XPGC_Rad.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRAUSUPER.

DR SMART; SM00279; HH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 KW Endonuclease.
 SQ SEQUENCE 382 AA; 42662 MW; 3A91LF83DC5B5A59 CRC64;
 Query Match 55.1%; Score 1068; DB 13; Length 382;
 Best Local Similarity 55.4%; Pred. No. 1.6e-67;
 Matches 209; Conservative 64; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGKGLTKLLADNAPKAMKEQKFSYFGKRIYAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
 DB 1 MGHLGLAKLIADVAPAAIKHEHDIKSYFGKRVAVDASMSIYQFLIVVGRGTGNETLTNEAGE 59
 QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
 DB 60 TTSHLGMGFYRTIRVHEGIRKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 119
 QY 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMGVPVVEAPSEAEAEACALCINDKVFVASED 180
 DB 120 VENIEKTRKLVKVTQHNEDCKRLRLMGVPVVEAPSEAEAEACALCINDKVFVASED 179
 QY 181 MSLIFGAPRFLRLHMDSPSKIPVMEFDVAKVLELELTMDQFIDLCILCGCDYCDSTK 240
 DB 180 MDALTFTGTVLLRHHTASEAKKLPIQEFHLNRVTDIGITHEQFVDFDLCILGSDYCTIR 239
 QY 241 GIGGQATKLIROHGSIESILENKNDRYQIPEDWPYQEARRLFKEPNVT-LDPELKW 299
 DB 240 GIGPKRAIDLROHKTIDEITIDNLDKYPVPENWLHKEAKHLFLEPEVVDITELKWI 299
 QY 300 APDEGLVAFNGEKFQFSEDRIRNGAKLAKNRQSTQGRLDFFKVTGSVSS-TKRKEA 358
 DB 300 EPDEGLVAFNGEKFQFSEDRIRNGAKLAKNRQSTQGRLDFFKVTGSVSS-TKRKEA 358
 QY 360 SDKTSKAAANKTKAGG 376
 DB 359 E---SKGSAKKAKTGG 372

RESULT 5
 P70054 PRELIMINARY; PRT; 382 AA.
 ID P70054
 AC P70054
 DT 01-FEB-1997 (TRENBLrel. 02, Created)
 DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE XEN1b.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEAD;
 RX MEDLINE=99069415; PubMed=9852084;
 RA Bibikova M., Wu B., Chi E., Kim K.H., Trautman J.K., Carroll D.,
 RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role
 in DNA metabolism".
 RL J. Biol. Chem. 273:34222-34229(1998).
 DR EMBL; U68141; AAB08478.1;
 DR HSSP; Q58839; 1A76
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N.1.
 DR InterPro; IPR003584; HHH_2.
 DR InterPro; IPR001532; XPGC_Rad.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRAUSUPER.
 DR SMART; SM00279; HH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.


```

SQ SEQUENCE 382 AA; 42865 MW; 1C648936A232D460 CRC64;
Query Match 55.1%; Score 1068; DB 13; Length 382;
Best Local Similarity 56.2%; Pred. No. 1.6e-67;
Matches 212; Conservative 60; Mismatches 99; Indels 6; Gaps 4;

QY 1 MGIGLTKLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
DB 1 MGIGHGLAKLIAERAPSARENDIKSYFGRKVAIDASMSIYQFLIAV-RQGAEVLQNEAGE 59

QY 61 VTSHQMGFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTAEVVG 120
DB 60 TTSHLGMGFYRTIRLMEHGKIPVYVFDGKPPDMKSGELAKRSERAEAKLEAAEEAGE 119

QY 121 KDAIEKLSKRTVKTROHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 180
DB 120 VENIEKFNKRLVKTROHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 179

QY 181 MDSLTGAPFLRLHMDPSSKKIPVMEFDVAKVLEBELTMDQFIDLCILGCGDYCDSIK 240
DB 180 MDALTFGTPVLLRLHMDPSSKKIPVMEFDVAKVLEBELTMDQFIDLCILGCGDYCETIR 239

QY 241 GIGGOTALKLIRHOGSIESILENLNDRYOIPEDWPYQEARLKEPNV-TLIDPELKW 299
DB 240 GIGPKRAIDLIRHOGSIESILENLNDRYOIPEDWPYQEARLKEPNV-TLIDPELKW 299

QY 300 APDEGLISFLVNDGNEFDVTKAEIKSAKNKSSQGRLESEFFKPTAT-TSAPLKRKE 358
DB 300 EPDEGLVAFMCGEKFQSEDIRNGAKLAKNRQSGTOGRLLDDFFKVTGTSISS-TKRKEY 358

QY 360 SDKTSKAAANKTKRAGG 376
DB 359 E---SKGTSKKKSTGG 372

RESULT 5
Q90YB0 PRELIMINARY; PRT; 381 AA.
AC Q90YB0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE FEN-1 nuclease.
GN FEN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsuzaki Y., Adachi N., Koyama H.;
RT "The FEN-1 nuclease is not essential for DNA replication, but is
RT required for maintenance of genomic integrity and base excision repair
RT in vertebrate cells."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB058602; BAB68507.1; -.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_1.
DR InterPro; IPR003583; HHH_1.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_1; 1.
DR SMART; SM00278; HHH1; 1.
DR PROSITE; PS00841; XPG_1; UNKNOWN_1.
DR PROSITE; PS00842; XPG_2; UNKNOWN_1.
KW Phosphopantetheine.
SQ SEQUENCE 381 AA; 43054 MW; 71432ECBCB4E67AA CRC64;
Query Match 54.8%; Score 1063.5; DB 13; Length 381;
Best Local Similarity 54.2%; Pred. No. 3.3e-67;
Matches 206; Conservative 67; Mismatches 104; Indels 3; Gaps 3;

```

```

QY 1 MGIGLTKLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
DB 1 MGIGHGLAKLIAERAPSARENDIKSYFGRKVAIDASMSIYQFLIAV-RQGAEVLQNEAGE 59

QY 61 VTSHQMGFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTAEVVG 120
DB 60 TTSHLGMGFYRTIRLMEHGKIPVYVFDGKPPDMKSGELAKRSERAEAKLEAAEEAGE 119

QY 121 KDAIEKLSKRTVKTROHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 180
DB 120 ENATIEKFNKRLVKTROHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 179

QY 181 MDSLTGAPFLRLHMDPSSKKIPVMEFDVAKVLEBELTMDQFIDLCILGCGDYCDSIK 240
DB 180 MDCLTFGSPVLMRLHMDPSSKKIPVMEFDVAKVLEBELTMDQFIDLCILGCGDYCESIR 239

QY 241 GIGGOTALKLIRHOGSIESILENLNDRYOIPEDWPYQEARLKEPNV-TLIDPELKW 299
DB 240 GIGPKRAIDLIRHOGSIESILENLNDRYOIPEDWPYQEARLKEPNV-TLIDPELKW 299

QY 300 APDEGLISFLVNDGNEFDVTKAEIKSAKNKSSQGRLESEFFKPTAT-TSAPLKRKE 358
DB 300 EPNEEELVQPMCGEKFQSEDIRNGVGRKLSKRSQSGTOGRLLDDFFKVTGTSISAKRKEPE 359

QY 359 TSDKTSKAAANKTKRAGG 378
DB 360 TKSAAKKAATNSATAKFKK 379

RESULT 7
Q91250 PRELIMINARY; PRT; 380 AA.
AC Q91250;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to flag structure-specific endonuclease.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010203; AAH10203.1; -.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_1.
DR InterPro; IPR003583; HHH_1.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_1; 1.
DR Pfam; PF00752; XPG_N; 1.
DR SMART; SM00278; HHH1; 1.
DR PROSITE; PS00841; XPG_1; UNKNOWN_1.
DR PROSITE; PS00842; XPG_2; UNKNOWN_1.
KW Endonuclease.
SQ SEQUENCE 380 AA; 42623 MW; 1BE903288B46520D CRC64;
Query Match 53.8%; Score 1043; DB 11; Length 380;
Best Local Similarity 53.8%; Pred. No. 9.2e-66;
Matches 205; Conservative 69; Mismatches 99; Indels 8; Gaps 5;

QY 1 MGIGLTKLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
DB 1 MGIGHGLAKLIAERAPSARENDIKSYFGRKVAIDASMSIYQFLIAV-RQGGDVQLNEAGE 59

QY 61 VTSHQMGFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTAEVVG 120
DB 60 TTSHLGMGFYRTIRMEHGKIPVYVFDGKPPDMKSGELAKRSERAEAKLEAAEEAGE 119

QY 121 KDAIEKLSKRTVKTROHNECDCKRLRLMGVPVVEAPSEAECAALCINDKVFVASED 180

```

Db 120 EEEVEKFTKRLVKVTKQHNDECKHLLSMGIPYLDAPSAEASCAALAKAGKYVAATED 179
Qy 181 MSLFTGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCLGCGDYCSIK 240
Db 180 MCLTFGSPVLMRHUTASEAKKLPQIEFHLSRVQLGSLNQEQFVLDCLGSLGSDYCSIR 239
Qy 241 GIGGQATKLIHQHSIESILENLKNDROYIPEDWPYQEARLKEPNVTL--IPELKW 298
Db 240 GIGPKRAVDLIQKHKSIEIIVRLDPSKYPVPENWLKHAQQLFEPEV-LDPESVELKW 298
Qy 299 TAPDEGLISFLVKGNGFEDRVTKAEIKSKAKNSOGRLSEFPFKPTATTAPLKRKE 358
Db 299 SEPNEELVKFCGKEQFSEIRISGVKRLSKSQGSTOGLRDDFFKVTGSLSS-AKRKE 357
Qy 359 TSDKTSKAAANKTKAGGKK 379
Db 358 PE---PKGPAKKAKTGAGK 375

RESULT 8
Q8R069 PRELIMINARY; PRT; 380 AA.
AC Q8R069
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE Hypothetical 42.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.
BL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027295; AAF27295.1;
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 42609 MW; C8B38745358F65AF CRC64;

Query Match 53.7%; Score 1042; DB 11; Length 380;
Best Local Similarity 53.5%; Pred. No. 1.1e-65;
Matches 204; Conservative 70; Mismatches 99; Indels 8; Gaps 5;
Qy 1 MGIKGLTKLLADNAPKAMKEQKESYFGKRIAYDASMSIYQFLIIVVGTGMEITNEAGE 60
Db 1 MGTHGLAKLIADVAPSAIRENDIKSYFGKRVADASMSIYQFLIIV--RQGGDVLQNEEGE 59
Qy 61 VTSHLOGMFRITRLLEAGIKPVYFDGKPPDMKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 60 TTSLLMGMYRTIRNMENGAVPVYFDGKPPQLKSGELASERRAEAKQLQAQAGM 119
Qy 121 KDAIEKLSRTVKVTRQHNDECKRLRLMGVPVVEAPSAEAEACALCINDKVFVASED 180
Db 120 EEEVEKFTKRLVKVTKQHNDECKHLLSMGIPYLDAPSAEASCAALAKAGKYVAATED 179
Qy 181 MSLFTGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCLGCGDYCSIK 240
Db 180 MCLTFGSPVLMRHUTASEAKKLPQIEFHLSRVQLGSLNQEQFVLDCLGSLGSDYCSIR 239
Qy 241 GIGGQATKLIHQHSIESILENLKNDROYIPEDWPYQEARLKEPNVTL--IPELKW 298
Db 240 GIGPKRAVDLIQKHKSIEIIVRLDPSKYPVPENWLKHAQQLFEPEV-LDPESVELKW 298
Qy 299 TAPDEGLISFLVKGNGFEDRVTKAEIKSKAKNSOGRLSEFPFKPTATTAPLKRKE 358
Db 299 SEPNEELVKFCGKEQFSEIRISGVKRLSKSQGSTOGLRDDFFKVTGSLSS-AKRKE 357
Qy 359 TSDKTSKAAANKTKAGGKK 379
Db 358 PE---PKGPAKKAKTGAGK 375

RESULT 9
Q9JHW7 PRELIMINARY; PRT; 380 AA.
AC Q9JHW7
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Flag structure-specific endonuclease.
GN Fen-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Cao G., Yang S., Li M., Chen J.;
RT "Cloning and characterization of a rat DNA structure-specific
RT endonuclease (Fen-1)."; the EMBL/GenBank/DBJ databases.
RL Submitted (Jun-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281018; AAF81265.1;
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N1.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGT; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00842; XPG_2; 1.
DR Endonuclease.
KW
SQ SEQUENCE 380 AA; 42622 MW; 61698CE16F182136 CRC64;

Query Match 53.6%; Score 1040; DB 11; Length 380;
Best Local Similarity 52.6%; Pred. No. 1.5e-65;
Matches 200; Conservative 73; Mismatches 101; Indels 6; Gaps 4;
Qy 1 MGIKGLTKLLADNAPKAMKEQKESYFGKRIAYDASMSIYQFLIIVVGTGMEITNEAGE 60
Db 1 MGTHGLAKLIADVAPSAIRENDIKSYFGKRVADASMSIYQFLIIV--RQGGDVLQNEEGE 59
Qy 61 VTSHLOGMFRITRLLEAGIKPVYFDGKPPDMKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 60 TTSLLMGMYRTIRNMENGAVPVYFDGKPPQLKSAQLAKRSERRAEAKQLQAQAGM 119
Qy 121 KDAIEKLSRTVKVTRQHNDECKRLRLMGVPVVEAPSAEAEACALCINDKVFVASED 180
Db 120 EEEVEKFTKRLVKVTKQHNDECKHLLSMGIPYLDAPSAEASCAALAKAGKYVAATED 179
Qy 181 MSLFTGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCLGCGDYCSIK 240
Db 180 MCLTFGSPVLMRHUTASEAKKLPQIEFHLSRVQLGSLNQEQFVLDCLGSLGSDYCSIR 239
Qy 241 GIGGQATKLIHQHSIESILENLKNDROYIPEDWPYQEARLKEPNVTL--IPELKW 299
Db 240 GIGPKRAVDLIQKHKSIEIIVRLDPSKYPVPENWLKHAQQLFEPEVLPESVELKWS 299
Qy 300 APDEGLISFLVKGNGFEDRVTKAEIKSKAKNSOGRLSEFPFKPTATTAPLKRKE 359
Db 300 EPNEELVKFCGKEQFSEIRISGVKRLSKSQGSTOGLRDDFFKVTGSLSS-AKRKE 358
Qy 360 SDKTSKAAANKTKAGGKK 379
Db 359 E---PKGPAKKAKTGAGK 375
RESULT 10
Q9N3T2 PRELIMINARY; PRT; 382 AA.
ID Q9N3T2
AC Q9N3T2

```
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 42.5 kDa protein.
GN Y47G6A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Minx P., Graves T., Hawrysko C.;
RT "The sequence of C. elegans cosmid Y47G6A."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024791; AAF60853.1;
DR HSSP; Q58839; IA76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_1.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR PROSITE; PS00842; XPG_2; 1.
DR Endonuclease.
KW Hypothetical protein.
SQ SEQUENCE 382 AA; 42549 MW; FDFE0BA0707321EB CRC64;

Query Match 53.5%; Score 1036.5; DB 5; Length 382;
Best Local Similarity 52.3%; Pred. No. 2.7e-65;
Matches 201; Conservative 73; Mismatches 99; Indels 11; Gaps 4;

QY 1 MGKIGTLKLLADNAPKAMKQKTESYFGKRIADVADSMISYQFLIVGRTGNETLTNEAGE 60
DB 1 MGKIGLSQVIADNAPSAIKYNEKMAFFGRTVAIDASMLYQFLIVADQDSQ-LQSEDEG 59
QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 60 TTSHLGMLNRTVMFENGKPVYVFDGKPPDMKKGELEKRSERRAEAKALTEAKEKG 119
QY 121 KDAIEKLSKRVKVTYRHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
DB 120 VKEAEKFERRLVKYTKQONDEAKRLGLMGIPVVEAPCEAEACALVKAGYEGTVD 179
QY 181 MDSLTFGAPRLRHLPDPSSKIPVMEFVAVKLELELTMDQFIDLCILGCDYCDISK 240
DB 180 MDAITFGSVTLRHLPAPVAKKIPKEFLNSLALAEKMLSVVEEFDLCILGCDYCGTIR 239
QY 241 GIGGQTALKLRQHGSTESILENLKNDRYQIPEDWPYQEARLKEPNVT-LDIPELKWT 299
DB 240 GVGPKAVELIRQHNKTIETILENDONKYPPEDPWPKRARELFNPEVTKRPEVELTWK 299
QY 300 APDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSQGRLESFF---KPTATTAPLKR 355
DB 300 EADVSGVIQFLGGEKNFNEIRNALKLKTSKSGTQGRIDSFFGNSTKVTCTVATKR 359
```

```
QY 356 RKETSDDKTSKAAANKTKAGGKK 379
DB 360 KAEAEAKAKGA----KGGPPK 378

RESULT 11
Q99M86 PRELIMINARY; PRT: 378 AA.
AC Q99M86;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Flap endonuclease-1.
GN FEN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV/J;
RA Karanjawala Z.E., Shi X., Hsieh C.-L., Lieber M.R.;
RT "The Mammalian Fen1 Locus: Structure and Conserved Sequence
RT Features."
RL Microb. Comp. Genomics 0:0-0(2001).
DR EMBL; AY014962; AAK01853.1;
DR HSSP; Q58839; IA76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N_1.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPGC_Rad.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00841; XPG_1; 1.
DR PROSITE; PS00842; XPG_2; 1.
DR Endonuclease.
KW Endonuclease.
SQ SEQUENCE 378 AA; 42300 MW; 827946BA8AC9F39 CRC64;

Query Match 52.0%; Score 1009; DB 11; Length 378;
Best Local Similarity 52.9%; Pred. No. 2.3e-63;
Matches 201; Conservative 69; Mismatches 102; Indels 8; Gaps 6;

QY 1 MGKIGTLKLLADNAPKAMKQKTESYFGKRIADVADSMISYQFLIVGRTGNETLTNEAGE 60
DB 1 MGKIGLAKLIADVAPSAIRENDINRTFGRKVAIDASMTYQFLIAY-RQGDVLYQNEEGE 59
QY 61 VTSHLQGMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
DB 60 TTS-LMGHFEYRTIR-MENGKIPVYVFDGKPPDLKSGELAKRSERRAEAKALQAOEAGM 117
QY 121 KDAIEKLSKRVKVTYRHNEDCKRLRLMGVPVVEAPSEAEACALCINDKVFVASED 180
DB 118 EBEVEKFKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALAKAKVYAAATED 177
QY 181 MDSLTFGAPRLRHLPDPSSKIPVMEFVAVKLELELTMDQFIDLCILGCDYCDISK 240
DB 178 MDCLTFGSPVLMRLHTASEAKLPIDQEFHLRVLEGLNGQFQVDCILGSDYCESIR 237
QY 241 GIGGQTALKLRQHGSTESILENLKNDRYQIPEDWPYQEARLKEPNVT-LDIPELKWT 299
DB 238 GIGAKRAYDLIRQHKISIEELVRDLDPKYPVFNMLHKEAQOLLEPEVDPSPSELKWS 297
QY 300 APDEGLISFLVKGNGFNEDRVTKAIEKISAKNKSQGRLESFFKPTATTAPLKRKET 359
DB 298 ENESELYKFMCGKQKQSEERISRGVRLSKRQSGTQGRLDDEFFNVTGSLSS-ARRKEP 356
QY 360 SDKTSKAAANKTKAGGKK 379
```

Db 357 E---PKGPAKKAKTGAGK 373

RESULT 12

O96830

ID O96830 PRELIMINARY; PRT; 385 AA.

AC

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

EG:EG0003.3 protein.

GN FEN1 OR EG:EG0003.3 OR CG8648.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

ON NCBI_TaxID=7227;

OX

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Abril J.F., Agayani A., An H.-J., Andrews-Pfannkoche C., Baldwin D.,

RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Fostler C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Glodek A., Hargreaves D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissentach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

[2]

RP SEQUENCE FROM N.A.

RA Murphy L., Harris D., Barrell B.;

RT "Sequencing the distal X chromosome of Drosophila melanogaster."

RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RA Benos P.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

DR ENBL; AE003805; AAF57944.1; -

DR ENBL; AL031863; CAA21320.1; -

DR HSP; Q58839; 1A76.

DR FlyBase; FBgn0225832; FEN1.

DR InterPro; IPR002421; 5_3_exonuclease.

DR InterPro; IPR000513; Exo_N_I.

DR InterPro; IPR003583; HHH_1.

DR InterPro; IPR003584; HHH_2.

DR InterPro; IPR001532; XPGC_Rad.

DR Pfam; PF01367; 5_3_exonuclease; 1.

DR Pfam; PF00867; XPG_I; 1.

DR Pfam; PF00752; XPG_N; 1.

DR PRINTS; PR00853; XPGRADSUPER.

DR SMART; SM00278; HHH1; 1.

DR SMART; SM00279; HHH2; 1.

DR SMART; SM00484; XPG1; 1.

DR SMART; SM00485; XPGN; 1.

DR PROSITE; PS00841; XPG_1; 1.

DR PROSITE; PS00842; XPG_2; 1.

SQ SEQUENCE 385 AA; 42948 MW; DOA0831C2BDA9240 CRC64;

Query Match 51.5%; Score 998.5; DB 5; Length 385;

Best Local Similarity 52.9%; Pred. No. 1.3e-62;

Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;

QY 1 MGIGKLTLLADNAPKAMKEQFESYFGRKIADVADSMISYQFLIVVGRGTGMETLTNEAGE 60

DB 1 MGILGLSKLIADLAPOAIRESEMKHFFGRKVAIDASMLYQFLIAVRSEGAQ-LATVNGD 59

QY 61 VTSHLQGMFNFTIRLLLEAGIKPVYFDGKPPDKKQKELAKRYSKRDDATKDLTEAVEVD 120

DB 60 PTHSLMGMYFTIRLLONGIKPVYFDGKPPDLKSGELAKRAERBEAEKALKAAATDAGD 119

QY 121 KDAIEKLSKRVKVTROHNECKRLRLMGVPVVEAPSEAEACALCINDKVFAYASED 180

DB 120 DAGIEKFNRLVRVTKHAKKELLMLGVPVVDAPCEAEACALVAKGVYATATED 179

QY 181 MDSLTFGAPRLHMLDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGDYCDISK 240

DB 180 MDALFGSTKLRLTYSEARKMPVKEFSDKLGLEINNNREFIDLCILCGDYCESIK 239

QY 241 GIGGOTALKLIHQHSIESILENINKDYOIPEDWPYQEARLRFEPNVT-LDIPELKWT 299

DB 240 GIGPRALELINTYRDIETILDNDSSKYTPVPMNWKVARELFIEPEVADADSIDLKWV 299

QY 300 APDEGLISFLVKONGFNEDRVTKAIFKISAKNKSOGLESFEK----PTATTSAPL 354

DB 300 EPDEGLVKFLCGDRQFNEVRNGAKLMKSKQAQTOVRDSEFFKTLSTPNATNA-- 357

QY 355 KRKETSDDTKAAANKTKAGG 376

DB 358 KRK--AEAAKKSANKKAKTSG 377

RESULT 13

O9GZ01 PRELIMINARY; PRT; 650 AA.

AC O9GZ01

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Flap endonuclease-1.

GN FEN1.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RA Casta L.J. III, Schmutte C., Taraschi T.F.;

RT "Flap Endonuclease-1 (Plasmodium falciparum).";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR ENBL; AF278764; AAG01445.1; -

DR HSP; Q58839; 1A76.

DR InterPro; IPR002421; 5_3_exonuclease.

DR InterPro; IPR000513; Exo_N_I.

DR InterPro; IPR003584; HHH_2.

DR InterPro; IPR001532; XPGC_Rad.

DR Pfam; PF01367; 5_3_exonuclease; 1.

```

DR Pfam: PF00867; XPG_I; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; Hh2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
KW Endonuclease.
SQ SEQUENCE 650 AA; 73989 MW; 2E9464DAB7F45B1D CRC64;

Query Match 45.7%; Score 887; DB 5; Length 650;
Best Local Similarity 46.6%; Pred. NO. 1.9e-54;
Matches 184; Conservative 74; Mismatches 117; Indels 20; Gaps 5;

QY 1 MGIKGLTKLLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVVGTGM-ETLTNEAG 59
DB 1 MGIKGLTKFTADAPNAIKIEKIESLMGRILAIADASMSLYQFIIRDSQYGNLTNESH 60

QY 60 EVTSHLQGFNRTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSKDDATKDLTEAVEVG 119
DB 61 ETTSHISGLMSIRLMEGLKPIYVFDGAPPELKGSELEKRGKRAEELKKAKEEG 120

QY 120 DKDAIEKLSKRTVKVTRHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASE 179
DB 121 NLEETIKKOSGRTVTRKONEEAKLLTLMGIPIIEAPCEAESQCAFLLKYNLAHATATE 180

QY 180 DMSLTGAPRFLRH------MDPSSKKIPVMEFDVAKVLELELTMDQFIDL 228
DB 181 DADALVFGTKILIRNLNANATSNQNNKNSKRGYILTEINLEQVLKGLNLTMDQFIDFC 240

QY 229 ILGCGDYCDTGKIGGQGTALKIRHGSIESILENKNDRYQIPEDWPYQEARLKEPN 288
DB 241 ILGCGDYCDTGKIGSKTAYNLKEYNCIEKIENIDQNKYQVPSFRFQEARKSFINPN 300

QY 289 VTLDIP---ELKWTAPDEEGLISFLVKGDFNEDRVTKAIEKISAKNKSOGRLSEPF 344
DB 301 V---LPKEDIKIDWNEPOIEELKHLKIDYNFNLRTVYNINLLKARKVTTQRRLDNEF 357

QY 345 KPTATTSAPLKKET-SDKTSKAAANKTKTAGKK 378
DB 358 TACTKSTKLIVETKKEQTLPAKGRKRTAGDK 392

RESULT 14
Q9U0K1
ID Q9U0K1 PRELIMINARY; PRT; 672 AA.
AC Q9U0K1.
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Flap endonuclease 1 (Putative flap exonuclease).
GN FEN1 OR MALP2.21.
OS Plasmodium falciparum, and
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833, 36329;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=P.falciparum; STRAIN=3D7A;
RA Li J.-L.;
RT "Characterization of the flap endonuclease 1 homolog from the human
RT malaria parasite Plasmodium falciparum.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=P.falciparum (isolate 3D7);
RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,
RA Quail M., Barrell B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF093702; AAG37989.1;
DR EMBL: AL035475; CAB62862.1;
DR HSP: Q58839; 1A76.
DR InterPro: IPR002421; 5_3_exonuclease.

```

```

DR InterPro: IPR000513; Exo_N_1.
DR InterPro: IPR003504; HHL_2.
DR InterPro: IPR001532; XPGC_Rad.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF00867; XPG_I; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; Hh2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
KW Endonuclease; Exonuclease.
SQ SEQUENCE 672 AA; 76680 MW; 6C2CDBB95F40042 CRC64;

Query Match 45.7%; Score 887; DB 5; Length 672;
Best Local Similarity 46.6%; Pred. NO. 2e-54;
Matches 184; Conservative 74; Mismatches 117; Indels 20; Gaps 5;

QY 1 MGIKGLTKLLADNAPKAMKEQKESYFGRKIAVDASMSIYQFLIVVGTGM-ETLTNEAG 59
DB 1 MGIKGLTKFTADAPNAIKIEKIESLMGRILAIADASMSLYQFIIRDSQYGNLTNESH 60

QY 60 EVTSHLQGFNRTIRLEAGIKPVYVFDGKPPDMKKOELAKRYSKDDATKDLTEAVEVG 119
DB 61 ETTSHISGLMSIRLMEGLKPIYVFDGAPPELKGSELEKRGKRAEELKKAKEEG 120

QY 120 DKDAIEKLSKRTVKVTRHNEDECKRLRLMGVPVVEAPSEAEACALCINDKVFVASE 179
DB 121 NLEETIKKOSGRTVTRKONEEAKLLTLMGIPIIEAPCEAESQCAFLLKYNLAHATATE 180

QY 180 DMSLTGAPRFLRH------MDPSSKKIPVMEFDVAKVLELELTMDQFIDL 228
DB 181 DADALVFGTKILIRNLNANATSNQNNKNSKRGYILTEINLEQVLKGLNLTMDQFIDFC 240

QY 229 ILGCGDYCDTGKIGGQGTALKIRHGSIESILENKNDRYQIPEDWPYQEARLKEPN 288
DB 241 ILGCGDYCDTGKIGSKTAYNLKEYNCIEKIENIDQNKYQVPSFRFQEARKSFINPN 300

QY 289 VTLDIP---ELKWTAPDEEGLISFLVKGDFNEDRVTKAIEKISAKNKSOGRLSEPF 344
DB 301 V---LPKEDIKIDWNEPOIEELKHLKIDYNFNLRTVYNINLLKARKVTTQRRLDNEF 357

QY 345 KPTATTSAPLKKET-SDKTSKAAANKTKTAGKK 378
DB 358 TACTKSTKLIVETKKEQTLPAKGRKRTAGDK 392

RESULT 15
Q8SS91
ID Q8SS91 PRELIMINARY; PRT; 345 AA.
AC Q8SS91.
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Structure-specific endonuclease of the XPG/RAD2 family.
GN ECU03_1080.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyretillade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite

```

```
RT Encephalitozoon cuniculi.*;
RL Nature 414:450-453(2001).;
DR EMBL: AL590433; CAD26252.1; -;
SQ SEQUENCE 345 AA; 38943 MW; B8C57D5F4B3E4F7C CRC64;

Query Match 43.48; Score 841.5; DB 5; Length 345;
Best Local Similarity 48.98; Pred. No. 1.4e-51;
Matches 170; Conservative 70; Mismatches 95; Indels 13; Gaps 5;

QY 1 MGKIGUTKLLADNAPKAMKEQKFESYFGRKIADVADSMISYQFLIVVGRGTGMTLTNEAGE 60
Db 1 MGKIQLSKLLRENSKRGIRERPLVYSSKKVAIDASMSYQFLIIV-RSGGATLGN-DS 58
QY 61 VTSHLQCMFNRTIRLLEAGIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 59 PTHLVGFFYRTIRMVDELGITPVYVFDGVPPEIKMKELEKERRAAADREVREASEVGD 118
QY 121 KDAIEKLSKRTVKVTRHNEPCKRLLRMGVVPVVEAPSEAECAALCINDKVFVASED 180
Db 119 KELMEYDRKTKVTGVHVECKRLLGLMGIPETAPSEAEAYCALLCKKXVYGVATED 178
QY 181 MDSITGAPFLEHLMDSPPSKIPVMEFDYAKVLELELTMDQFIDLCILCGCDYCDISIK 240
Db 179 MDALTEGSPVLRNFGTQSRRLPVMENLPQILEDLSLDHSEFIDLCILCGCDYCSYTK 238
QY 241 GIGQOTALKLIRQHSIESILENLKNDYQIPEDWYQFARRLFKEPNVYTLDIPELK--- 297
Db 239 GIGPKKALGLIKKRSIGNILKN---EDLEVPGDWRYSDAQKIF---GSLAEIGEIRDFN 292
QY 298 --WTAPDEGLISFLVKDNGFNEDRVTKAEIKKSARKNSQGRLESF 343
Db 293 ISWDSIDRNGIVNFLVEEKGFDLERNVKGIDKLINSKKKGTOGRLDGF 340
```

Search completed: May 8, 2003, 21:42:47
Job time : 61 secs

